



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C07H 21/02, 21/04, C12N 1/20, 15/00, 15/09, 15/63, 15/70, 15/74</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 00/61596</b> <b>(43) International Publication Date:</b> 19 October 2000 (19.10.00)
<b>(21) International Application Number:</b> PCT/US00/08983 <b>(22) International Filing Date:</b> 6 April 2000 (06.04.00)  <b>(30) Priority Data:</b> 60/128,703 9 April 1999 (09.04.99) US 60/176,068 20 January 2000 (20.01.00) US  <b>(71) Applicants (for all designated States except US):</b> HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850 (US). ROSEN, Craig, A. [US/US]; 22400 Rolling Hill Road, Laytonsville, MD 20882 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> RUBEN, Steven, M. [US/US]; 18528 Heritage Hills Drive, Laytonsville, MD 20882 (US). KOMATSOULIS, George [US/US]; 9518 Garwood Street, Silver Spring, MD 20901 (US).  <b>(74) Agents:</b> HOOVER, Kenley, K. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).		<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> 50 HUMAN SECRETED PROTEINS  <b>(57) Abstract</b>  The present invention relates to novel human secreted proteins and isolated nucleic acids containing the coding regions of the genes encoding such proteins. Also provided are vectors, host cells, antibodies, and recombinant methods for producing human secreted proteins. The invention further relates to diagnostic and therapeutic methods useful for diagnosing and treating diseases, disorders, and/or conditions related to these novel human secreted proteins.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						



## 50 Human Secreted Proteins

### *Field of the Invention*

This invention relates to newly identified polynucleotides and the polypeptides encoded by these polynucleotides, uses of such polynucleotides and polypeptides, and their production.

### *Background of the Invention*

Unlike bacterium, which exist as a single compartment surrounded by a membrane, human cells and other eucaryotes are subdivided by membranes into many functionally distinct compartments. Each membrane-bounded compartment, or organelle, contains different proteins essential for the function of the organelle. The cell uses "sorting signals," which are amino acid motifs located within the protein, to target proteins to particular cellular organelles.

One type of sorting signal, called a signal sequence, a signal peptide, or a leader sequence, directs a class of proteins to an organelle called the endoplasmic reticulum (ER). The ER separates the membrane-bounded proteins from all other types of proteins. Once localized to the ER, both groups of proteins can be further directed to another organelle called the Golgi apparatus. Here, the Golgi distributes the proteins to vesicles, including secretory vesicles, the cell membrane, lysosomes, and the other organelles.

Proteins targeted to the ER by a signal sequence can be released into the extracellular space as a secreted protein. For example, vesicles containing secreted proteins can fuse with the cell membrane and release their contents into the extracellular space - a process called exocytosis. Exocytosis can occur constitutively or after receipt of a triggering signal. In the latter case, the proteins are stored in secretory vesicles (or secretory granules) until exocytosis is triggered. Similarly,

proteins residing on the cell membrane can also be secreted into the extracellular space by proteolytic cleavage of a "linker" holding the protein to the membrane.

Despite the great progress made in recent years, only a small number of genes encoding human secreted proteins have been identified. These secreted proteins  
5 include the commercially valuable human insulin, interferon, Factor VIII, human growth hormone, tissue plasminogen activator, and erythropoietin. Thus, in light of the pervasive role of secreted proteins in human physiology, a need exists for identifying and characterizing novel human secreted proteins and the genes that encode them. This knowledge will allow one to detect, to treat, and to prevent  
10 medical diseases, disorders, and/or conditions by using secreted proteins or the genes that encode them.

### ***Summary of the Invention***

The present invention relates to novel polynucleotides and the encoded  
15 polypeptides. Moreover, the present invention relates to vectors, host cells, antibodies, and recombinant and synthetic methods for producing the polypeptides and polynucleotides. Also provided are diagnostic methods for detecting diseases, disorders, and/or conditions related to the polypeptides and polynucleotides, and therapeutic methods for treating such diseases, disorders, and/or conditions. The  
20 invention further relates to screening methods for identifying binding partners of the polypeptides.

### ***Detailed Description***

#### **Definitions**

25 The following definitions are provided to facilitate understanding of certain terms used throughout this specification.

In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be  
5 contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA  
10 preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide/sequences of the present invention.

In the present invention, a "secreted" protein refers to those proteins capable of being directed to the ER, secretory vesicles, or the extracellular space as a result of  
15 a signal sequence, as well as those proteins released into the extracellular space without necessarily containing a signal sequence. If the secreted protein is released into the extracellular space, the secreted protein can undergo extracellular processing to produce a "mature" protein. Release into the extracellular space can occur by many mechanisms, including exocytosis and proteolytic cleavage.

20 In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as  
25 disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding

sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

5           As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:X or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without the signal sequence, the  
10   secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

          In the present invention, the full length sequence identified as SEQ ID NO:X  
15   was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:X was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard,  
20   Manassas, Virginia 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure.

          A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to  
25   sequences contained in SEQ ID NO:X, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an

overnight incubation at 42 degree C in a solution comprising 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20  $\mu$ g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65 degree C.

5 Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or  
10 temperature. For example, lower stringency conditions include an overnight incubation at 37 degree C in a solution comprising 6X SSPE (20X SSPE = 3M NaCl; 0.2M  $\text{NaH}_2\text{PO}_4$ ; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100  $\mu$ g/ml salmon sperm blocking DNA; followed by washes at 50 degree C with 1XSSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed  
15 following stringent hybridization can be done at higher salt concentrations (e.g. 5X SSC).

Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include  
20 Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

Of course, a polynucleotide which hybridizes only to polyA+ sequences (such  
25 as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a

complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

5       The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of  
10 single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA  
15 backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

      The polypeptide of the present invention can be composed of amino acids  
20 joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more  
25 detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid

side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992).)

"SEQ ID NO:X" refers to a polynucleotide sequence while "SEQ ID NO:Y" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

"A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the

present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not  
5 more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention.)

Many proteins (and translated DNA sequences) contain regions where the amino acid composition is highly biased toward a small subset of the available  
10 residues. For example, membrane spanning domains and signal peptides (which are also membrane spanning) typically contain long stretches where Leucine (L), Valine (V), Alanine (A), and Isoleucine (I) predominate. Poly-Adenosine tracts (polyA) at the end of cDNAs appear in forward translations as poly-Lysine (poly-K) and poly-Phenylalanine (poly-F) when the reverse complement is translated. These regions are  
15 often referred to as “low complexity” regions.

Such regions can cause database similarity search programs such as BLAST to find high-scoring sequence matches that do not imply true homology. The problem is exacerbated by the fact that most weight matrices (used to score the alignments  
20 generated by BLAST) give a match between any of a group of hydrophobic amino acids (L,V and I) that are commonly found in certain low complexity regions almost as high a score as for exact matches.

In order to compensate for this, BLASTX.2 (version 2.0a5MP-WashU) employs two filters (“seg” and “xnu”) which “mask” the low complexity regions in a  
25 particular sequence. These filters parse the sequence for such regions, and create a new sequence in which the amino acids in the low complexity region have been



replaced with the character "X". This is then used as the input sequence (sometimes referred to herein as "Query" and/or "Q") to the BLASTX program. While this regime helps to ensure that high-scoring matches represent true homology, there is a negative consequence in that the BLASTX program uses the query sequence that has  
5 been masked by the filters to draw alignments.

Thus, a stretch of "X"s in an alignment shown in the following application does not necessarily indicate that either the underlying DNA sequence or the translated protein sequence is unknown or uncertain. Nor is the presence of such stretches meant to indicate that the sequence is identical or not identical to the  
10 sequence disclosed in the alignment of the present invention. Such stretches may simply indicate that the BLASTX program masked amino acids in that region due to the detection of a low complexity region, as defined above. In all cases, the reference sequence(s) (sometimes referred to herein as "Subject", "Sbjct", and/or "S") indicated in the specification, sequence table (Table 1), and/or the deposited clone is (are) the  
15 definitive embodiment(s) of the present invention, and should not be construed as limiting the present invention to the partial sequence shown in an alignment, unless specifically noted otherwise herein.

### **Polynucleotides and Polypeptides of the Invention**

20

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 1**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: breast lymph node CDNA library and to a lesser extent in Brain frontal cortex; Breast Lymph node cDNA library; HEL cell line; Prostate BPH;  
25 Human umbilical vein endothelial cells, IL-4 induced; Soares breast 2NbHBst;

Human Substantia Nigra; NCI\_CGAP\_Lip2; Soares\_testis\_NHT and Hodgkin's Lymphoma II.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:11 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1794 of SEQ ID NO:11, b is an integer of 15 to 1808, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:11, and where b is greater than or equal to a + 14.

## 15 FEATURES OF PROTEIN ENCODED BY GENE NO: 2

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil3043919 (all information available through the recited accession number is incorporated herein by reference, for example, J. Biol. Chem. 273 (9), 4827-4830 (1998), and Immunity 9, 13-24 (1998)) which is described therein as "docking protein [Homo sapiens]". A partial alignment demonstrating the observed homology is shown immediately below.

25 >gi|3043919 (AF034970) docking protein [Homo sapiens] >sp|O60496|O60496  
DOCKING

PROTEIN.  
Length = 412

30 Plus Strand HSPs:

Score = 1445 (508.7 bits), Expect = 2.7e-147, P = 2.7e-147  
 Identities = 302/412 (73%), Positives = 324/412 (78%), Frame = +1

```

5   Q:   70 MGDGAVKQGFLYLQQQQTFGKKWRRFGASLYGGSDCALARLELQEGPEKPRRCEAARKVI 249
      MGDGAVKQGFLYLQQQQTFGKKWRRFGASLYGGSDCALARLELQEGPEKPRRCEAARKVI
      S:   1 MGDGAVKQGFLYLQQQQTFGKKWRRFGASLYGGSDCALARLELQEGPEKPRRCEAARKVI 60

10  Q:  250 RLSDCLRVAEAGGEASSPRDTSAFFLETKERLYLLAAPAAERGDWVQAICLLAFPGQRKE 429
      RLSDCLRVAEAGGEASSPRDTSAFFLETKERLYLLAAPAAERGDWVQAICLLAFPGQRKE
      S:   61 RLSDCLRVAEAGGEASSPRDTSAFFLETKERLYLLAAPAAERGDWVQAICLLAFPGQRKE 120

      Q:  430 LSGPEGKQSRPFMGKSNCTACAVTVGPHKEFAVTMRPTEASERCHLRGSYTLRAGESALE 609
      LSGPEGKQSRP M ++ + AVTVGPHKEF VTMRPTEASERCH RGSYTLRAGES+LE
15  S:  121 LSGPEGKQSRPCMEENELYSSAVTVGPHKEFPVTMRPTEASERCHRRGSYTLRAGESSE 180

      Q:  610 LWGGPKPGTQLYDWPYRFLRRFGRDKAVAASLE--RATLSSKPGKAMRSSWPWKRPSPR 783
      LWGGP+PGTQLYDWPYRFLRRFGRDK V S E R +S + G + L
      S:  181 LWGGPEPGTQLYDWPYRFLRRFGRDK-VTFSFEAGRRCVSGE-GNFEFETRQNEIFLAL 238

20  Q:  784 RMLHPLHPNRSQPQSPPRCPGLIAPTLGRMTHCRR LHPPHRCLLHGLGARRGSMPCPSMR 963
      N + P +P P I +L R PH L + +P P R
      S:  239 EEAISAKNAA-PATPQPQATIPASLPRPDS--PYSRPHDSLPPP--SPTTPVPAPRPR 293

25  Q:  964 -----WPVPW---GRTSGASWQSLSSWPTLL----YDSIEETLPPRPDHIYDEPEGVAA 1107
      + VP+ R+ G +++ +L+ P LL YDSIEETLPPRPDHIYD+PEGVAA
      S:  294 GQEGEYAVPFDVARSLGKNFRGILAVPPQLLADPLYDSIEETLPPRPDHIYDKPEGVAA 353

30  Q: 1108 LSLYDSPQEPRGEAWRRQATADRDPAQLQHVQPAGQDFSASGWQPGTEYDNVVLKKGPK 1284
      LSLYDSPQEPRGEAWRRQAT DRDPAGLQHVQPAGQDFSASGWQPGTEYDNVVLKKGPK
      S:  354 LSLYDSPQEPRGEAWRRQATDRDPAGLQHVQPAGQDFSASGWQPGTEYDNVVLKKGPK 412

```

The segment of gil3043919 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 111. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 112 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Activated T-cells, 24 hrs, re-excision and to a lesser extent in Activated T-Cells, 12 hrs, re-excision; Activated T-Cells, 12 hrs.; Activated T-Cells, 24 hrs.; Primary Dendritic Cells, lib 1; Soares\_pregnant\_uterus\_NbHPU; Human Adult Pulmonary, re-excision; Activated T-Cell (12hs)/Thiouridine labelled Eco; Activated T-cell(12h)/Thiouridine-re-excision; Activated T-Cells, 8 hrs.; Human Tonsil, Lib 3; eosinophil-IL5 induced; Human Adipose Tissue, re-excision; HL-60, PMA 4H, re-excision; Human Osteosarcoma; Human Activated T-Cells; Human Adipose; Ulcerative Colitis; Human Eosinophils; Spleen, Chronic lymphocytic leukemia; Neutrophils IL-1 and LPS induced and T cell helper II.

The gene encoding the disclosed cDNA is believed to reside on chromosome 8. Accordingly, polynucleotides related to this invention are useful as a marker in linkage analysis for chromosome 8.

The tissue distribution in activated T-cells, combined with the homology to the docking protein suggests the protein product of this clone is useful for the diagnosis and treatment of a variety of immune system disorders, particularly Chronic myelogenous leukemia (CML). Moreover, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).

Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease,

inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue  
5 injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the  
10 expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. The uses include bone marrow cell ex- vivo culture, bone marrow transplantation, bone marrow reconstitution, radiotherapy or chemotherapy of neoplasia. The gene product may also be involved in lymphopoiesis, therefore, it can be used in immune disorders such as  
15 infection, inflammation, allergy, immunodeficiency etc. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed  
20 tissues.

The secreted protein can also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, and as nutritional supplements. It may also have a very wide range of biological activities. Typical of these are cytokine, cell  
25 proliferation/differentiation modulating activity or induction of other cytokines; immunostimulating/immunosuppressant activities (e.g. for treating human

immunodeficiency virus infection, cancer, autoimmune diseases and allergy);  
regulation of hematopoiesis (e.g. for treating anemia or as adjunct to chemotherapy);  
stimulation or growth of bone, cartilage, tendons, ligaments and/or nerves (e.g. for  
treating wounds, stimulation of follicle stimulating hormone (for control of fertility);  
5 chemotactic and chemokinetic activities (e.g. for treating infections, tumors);  
hemostatic or thrombolytic activity (e.g. for treating hemophilia, cardiac infarction  
etc.); anti-inflammatory activity (e.g. for treating septic shock, Crohn's disease); as  
antimicrobials; for treating psoriasis or other hyperproliferative diseases; for  
regulation of metabolism, and behavior. Also contemplated is the use of the  
10 corresponding nucleic acid in gene therapy procedures.

Similarly, developmental tissues rely on decisions involving cell  
differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can  
result in inappropriate suppression of cell death, as occurs in the development of some  
cancers, or in failure to control the extent of cell death, as is believed to occur in  
15 acquired immunodeficiency and certain neurodegenerative disorders, such as spinal  
muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the  
present invention are useful in treating, detecting, and/or preventing said disorders  
and conditions, in addition to other types of degenerative conditions. Thus this protein  
may modulate apoptosis or tissue differentiation and would be useful in the detection,  
20 treatment, and/or prevention of degenerative or proliferative conditions and diseases.  
The protein is useful in modulating the immune response to aberrant polypeptides, as  
may exist in proliferating and cancerous cells and tissues. The protein can also be  
used to gain new insight into the regulation of cellular growth and proliferation.  
Protein, as well as, antibodies directed against the protein may show utility as a tumor  
25 marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:12 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1729 of SEQ ID NO:12, b is an integer of 15 to 1743, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:12, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 3

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil1710216 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "Homo sapiens protein". A partial alignment demonstrating the observed homology is shown immediately below.

```
>gi|1710216 unknown [Homo sapiens]
      Length = 139
```

Plus Strand HSPs:

```
Score = 160 (56.3 bits), Expect = 3.1e-11, P = 3.1e-11
Identities = 35/56 (62%), Positives = 42/56 (75%), Frame = +1
```

```
Q:      1 ETESCSVS--ECSGMVSVHCNHLHLLGSSDSPASVSQVAGITGTCQHAQL*TIF*AE 162
      ET+S SV+  ECSG +S HCNL L GSS+SPAS S+VAG  GTC+ AQL  +F AE
S:      54 ETQSHSVTRLECSGTISAHCNLCPLPGSSNSPASASRVAGTAGTCRRAQLIFVFLAE 109
```

The segment of gil1710216 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 113. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein.

5 Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 114 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
10 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares fetal liver spleen 1NFLS and to a lesser extent in Soares placenta Nb2HP; Soares melanocyte 2NbHM; Human Fetal Kidney; Soares adult brain N2b5HB55Y; Early Stage Human Brain; Human Fetal Lung III;  
15 Soares\_senescent\_fibroblasts\_NbHSF; HUMAN JURKAT MEMBRANE BOUND POLYSOMES; Human Uterine Cancer; Endothelial-induced; HUMAN B CELL LYMPHOMA; Human Cerebellum; Human Umbilical Vein, Endo. remake; Amniotic Cells - Primary Culture; Stromal cell TF274; Soares\_parathyroid\_tumor\_NbHPA; Human Greater Omentum, fract II remake,; Smooth Muscle- HASTE normalized;  
20 Human Lung Cancer, re-excision; Human Stomach, re-excision; Human endometrial stromal cells-treated with progesterone; Ovarian Tumor 10-3-95; Human Osteosarcoma; Stratagene HeLa cell s3 937216; Stratagene colon (#937204); Breast Cancer Cell line, angiogenic; Macrophage-oxLDL; Liver, Hepatoma; Macrophage (GM-CSF treated); Macrophage-oxLDL, re-excision; Smooth muscle, serum treated;  
25 H Macrophage (GM-CSF treated), re-excision; Normal colon; Activated T-Cell (12hs)/Thiouridine labelledEco; Endothelial cells-control; Human Microvascular



Endothelial Cells, fract. A; Monocyte activated; Bone Marrow Cell Line (RS4,11);  
 Keratinocyte; Soares ovary tumor NbHOT; Smooth Muscle Serum Treated, Norm;  
 Amniotic Cells - TNF induced; NTERA2 teratocarcinoma cell line+retinoic acid (14  
 days); Raji Cells, cyclohexamide treated; Human Normal Breast; Human endometrial  
 5 stromal cells-treated with estradiol; Human Fetal Epithelium (Skin); Human Adipose  
 Tissue, re-excision; HL-60, PMA 4H, re-excision; Jurkat T-Cell, S phase; KMH2;  
 Fetal Liver, subtraction II; Apoptotic T-cell; Soares\_fetal\_heart\_NbHH19W; Smooth  
 muscle, serum induced, re-exc; 12 Week Old Early Stage Human; Colon Carcinoma;  
 Soares breast 3NbHBst; Human Testes Tumor; Bone marrow; Spleen, Chronic  
 10 lymphocytic leukemia; Raji cells, cyclohexamide treated, differentially expressed;  
 Human Thymus Tumor, subtracted; LNCAP + 0.3nM R1881; Human Aortic  
 Endothelium; Human Cerebellum, subtracted; Soares\_fetal\_heart\_NbHH19W; Supt  
 Cells, cyclohexamide treated; Activated T-cells; Human Thyroid; H. cerebellum,  
 Enzyme subtracted; Human Liver; Activated T-Cells, 12 hrs, re-excision; Cem cells  
 15 cyclohexamide treated; Mesangial cell, frac 2; Hepatocellular Tumor; Smooth  
 muscle, IL1b induced; Breast Cancer cell line, MDA 36; Healing groin wound, 7.5  
 hours post incision; Stratagene endothelial cell 937223; Stratagene hNT neuron  
 (#937233); Stratagene endothelial cell 937223; Human endometrial stromal cells;  
 Jurkat T-cell G1 phase; Stratagene lung carcinoma 937218; Myeloid Progenitor Cell  
 20 Line; Soares\_fetal\_heart\_NbHH19W; Human Pituitary, subt IX; H. Kidney Medulla,  
 re-excision; Human Thymus; Human Dermal Endothelial Cells, untreated; L428;  
 Stratagene fetal spleen (#937205); T-Cell PHA 24 hrs; Human Activated Monocytes;  
 Human Rhabdomyosarcoma; Synovial Fibroblasts (control); Human Thymus;  
 Hemangiopericytoma; Human Fetal Brain; Human Gall Bladder; Resting T-Cell  
 25 Library, II; Soares\_fetal\_lung\_NbHL19W; Human Fetal Heart; Colon Normal III;  
 Anergic T-cell; Human Amygdala; Smooth muscle, control;

Soares\_fetal\_heart\_NbHH19W; Stratagene endothelial cell 937223; Human Testes;  
Nine Week Old Early Stage Human; Soares\_fetal\_lung\_NbHL19W;  
Soares\_placenta\_8to9weeks\_2NbHP8to9W; Soares infant brain 1NIB; Activated T-  
Cells, 4 hrs., ligation 2; Activated T-Cells, 4 hrs.; Activated T-Cells, 24 hrs., ligation  
5 2; Human Thymus, subtracted; HCC cell line metastasis to liver; LNCAP, differential  
expression; Human Fetal Kidney; Human Fetal Brain, normalized CO; Human  
Eosinophils; Human Macrophage; Human Adult Spleen, fractionII; L428 cell line;  
Human Macrophage, subtracted; Human Infant Adrenal Gland, Subtracted; Palate  
carcinoma; Larynx Carcinoma; Human Kidney Cortex, re-rescue; Human Fetal Liver,  
10 subtracted, neg clone; H. Adipose Tissue; Human Uterus, normal; Human Prostate  
Cancer, Stage B2 fraction; H. Normalized Fetal Liver, II; Activated T-Cells, 12 hrs.;  
Dermatofibrosarcoma Protuberance; Human Adult Liver, subtracted; Human  
Pituitary, re-excision; Human Adult Spleen; Human Fetal Brain; Human (HCC) cell  
line liver (mouse) metastasis, remake; Human Colon Carcinoma (HCC) cell line;  
15 Human Adult Retina; Human Umbilical Vein Endothelial Cells, fract. A; Human  
Pancreatic Carcinoma; Human Pituitary, subtracted; Frontal lobe,dementia,re-  
excision; Smooth muscle-ILb induced; Human Lung; Human Soleus; Human Adult  
Heart,re-excision; Resting T-Cell, re-excision; Human Pineal Gland; Apoptotic T-  
cell, re-excision; HEL cell line; Stomach cancer (human),re-excision; Human  
20 Amygdala,re-excision; STROMAL -OSTEOCLASTOMA; Human Colon Cancer,re-  
excision; Human Prostate Cancer, Stage C fraction; Synovial IL-1/TNF stimulated;  
NTERA2 + retinoic acid, 14 days; H. Kidney Cortex, subtracted;  
Soares\_pregnant\_uterus\_NbHPU; Salivary Gland, Lib 2; Healing groin wound, 6.5  
hours post incision; Stratagene muscle 937209; Human Osteoclastoma, re-excision;  
25 Synovial hypoxia; Soares\_pregnant\_uterus\_NbHPU; Soares\_fetal\_lung\_NbHL19W;  
Soares\_parathyroid\_tumor\_NbHPA; Human Ovary; Human Manic Depression

Tissue; Human Chronic Synovitis; Human Prostate; T-Cell PHA 16 hrs; Human Brain, Striatum; human ovarian cancer; 12 Week Old Early Stage Human, II; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human Hypothalamus, Schizophrenia; Human umbilical vein endothelial cells, IL-4 induced; Spinal cord; Human

5 Chondrosarcoma; Epithelial-TNF $\alpha$  and INF induced; Bone Marrow Stromal Cell, untreated; Human Adrenal Gland Tumor; Rejected Kidney, lib 4; Human Whole Six Week Old Embryo; Stratagene lung (#937210); Human Liver, normal; Soares\_fetal\_lung\_NbHL19W; Colon Tumor; H. Frontal cortex, epileptic, re-excision; Human T-Cell Lymphoma; Soares\_parathyroid\_tumor\_NbHPA; Adipocytes;

10 Stratagene muscle 937209; Dendritic cells, pooled; Colon Tumor II; Soares\_fetal\_lung\_NbHL19W; Human Synovial Sarcoma; Human Neutrophil, Activated; Human Adult Pulmonary, re-excision; CD34 depleted Buffy Coat (Cord Blood), re-excision; T Cell helper I; Human Bone Marrow, treated; Soares\_multiple\_sclerosis\_2NbHMSP; Neutrophils IL-1 and LPS induced; Activated

15 T-cell(12h)/Thiouridine-re-excision; Human Endometrial Tumor; Hodgkin's Lymphoma II; neutrophils control; Soares\_fetal\_heart\_NbHH19W; Human 8 Week Whole Embryo; Pancreatic Islet; Soares\_senescent\_fibroblasts\_NbHSF and Primary Dendritic Cells, lib 1.

Many polynucleotide sequences, such as EST sequences, are publicly

20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:13 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention

25 are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 613 of SEQ ID NO:13, b

is an integer of 15 to 627, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:13, and where b is greater than or equal to a + 14.

## **5 FEATURES OF PROTEIN ENCODED BY GENE NO: 4**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human Thymus; Human Fetal Spleen; CD34 depleted Buffy Coat (Cord Blood), re-excision.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:14 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 521 of SEQ ID NO:14, b is an integer of 15 to 535, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:14, and where b is greater than or equal to a + 14.

20

## **FEATURES OF PROTEIN ENCODED BY GENE NO: 5**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human Testes and to a lesser extent in Human Testes, Reexcision; Human Umbilical Vein, Reexcision; Human Activated T-Cells; Human adult testis, large inserts; human tonsils; Human 8 Week Whole Embryo, subtracted; Human Osteoblasts II; Human umbilical vein endothelial cells, IL-4 induced;

25

NCI\_CGAP\_Thy1; Soares\_NFL\_T\_GBC\_S1; Hodgkin's Lymphoma II and Primary Dendritic Cells, lib 1.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 65 as residues: Arg-26 to Ser-37.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:15 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence  
10 would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 705 of SEQ ID NO:15, b is an integer of 15 to 719, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:15, and where b is greater than or equal to a  
15 + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 6

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting  
20 example, the sequence accessible through the following database accession no. gi|1710216 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "Homo sapiens protein". A partial alignment demonstrating the observed homology is shown immediately below.

25

```
>gi|1710216 unknown [Homo sapiens]  
Length = 139
```

Plus Strand HSPs:

Score = 157 (55.3 bits), Expect = 9.0e-23, Sum P(3) = 9.0e-23  
Identities = 31/45 (68%), Positives = 36/45 (80%), Frame = +3

5 Q: 57 RLECSGAVTAHCNLSLLGSSDPPASASRVAGTTGTHHHAWLIFLY 191  
RLECSG ++AHCNL L GSS+ PASASRVAGT GT A LIF++  
S: 62 RLECSGTISAHCNLCLPGSSNSPASASRVAGTAGTCRRAQLIFVF 106

10 Score = 83 (29.2 bits), Expect = 9.0e-23, Sum P(3) = 9.0e-23  
Identities = 14/19 (73%), Positives = 17/19 (89%), Frame = +2

15 Q: 185 FVFLVDMGFHHVSQDGLDV 241  
FVFL +MGFHHV +DGLD+  
S: 104 FVFLAEMGFHHVGRDGLDL 122

20 Score = 83 (29.2 bits), Expect = 0.54, P = 0.42  
Identities = 15/19 (78%), Positives = 17/19 (89%), Frame = +2

Q: 845 ECNGTISTHCNLCLPCGTS 901  
EC+GTIS HCNLCLP G+S  
S: 64 ECSGTISAHCNLCLP-GSS 81

The segments of gil1710216 that are shown as "S" above are set out in the sequence listing as SEQ ID NO. 115, SEQ ID NO. 117 and SEQ ID NO. 119. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 116, SEQ ID NO. 118 and/or SEQ ID NO. 120 which correspond to the "Q" sequences in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in Neutrophils IL-1 and LPS induced.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 66 as residues: Pro-40 to Tyr-46.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:16 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 887 of SEQ ID NO:16, b is an integer of 15 to 901, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:16, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 7**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares breast 3NbHBst and to a lesser extent in Human White Fat; Human Skin Tumor; Human Prostate Cancer, Stage C fraction; Human adult (K.Okubo); NCI\_CGAP\_Pr12; Healing groin wound, 7.5 hours post incision; LNCAP prostate cell line and Human Amygdala.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:17 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 905 of SEQ ID NO:17, b

is an integer of 15 to 919, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:17, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 8

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gnl|PID|e311150 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "putative 3,4-dihydroxy-2-butanone kinase [*Lycopersicon esculentum*]". A partial alignment demonstrating the observed homology is shown immediately below.

```

15  >gnl|PID|e311150 putative 3,4-dihydroxy-2-butanone kinase [Lycopersicon
      esculentum]
      Length = 594

      Plus Strand HSPs:

20  Score = 237 (83.4 bits), Expect = 1.8e-18, P = 1.8e-18
      Identities = 61/186 (32%), Positives = 89/186 (47%), Frame = +3

25  Q:   69 LERV CSTLLGLEEHLNALDRAAGD GDCG TTHSRAARAIQEWLKEGPPPASPAQXXXXXXXXX 248
      +E  + ++ L ++LN  D   GDGDCG+T  R A AI E +K+  P   PA+
      S:  387 IEAAATEVVNLRDNLNEWDNKVGDGDCGSTMFRGAVAILED MKKYYPLNDPAETVNEIGA 446

      Q:   249 XXXEKMGGSSGALYGLFLTAAQPLK--AKTSLPA--WSAAMDAGLEAMQKYGKAAPGDR 416
      MGG+SG LY +F  AA   LK  A++ + A  W+ A++A + A+ KYG A+ G R
      S:  447 SIGRVMGGTSGILYSIFCKAAYAKLKENAESVVTAIHWADALEAAIAAVSKYGGASAGYR 506

30  Q:   417 TMLDSLWAAGQELQAWKSPGADLLQVLTXXXXXXXXXXXXXXXXXXXXXGRASYISSARLE 596
      T+LD+L  A   L+   + G D                               GR++Y+   L
      S:  507 TLLDALIPALSALKERLNAGDDPADAFIISAEASAGAESTKHMQAQAGRSTYVPGDILA 566

35  Q:   597 Q-PDPG 611
      PDPG
      S:  567 SVPDPG 572

```



The segment of gnlPIDle311150 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 121. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein.

5 Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 122 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
10 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Supt Cells, cyclohexamide treated; Synovial hypoxia-RSF subtracted; Human Pituitary, subt IX and to a lesser extent in Human Pituitary, subtracted V; H. cerebellum, Enzyme subtracted; Resting T-Cell, re-excision; Human  
15 Prostate Cancer, Stage C fraction; Human Adult Small Intestine; Human Fetal Dura Mater; T-Cell PHA 24 hrs; Human Activated Monocytes; Human Fetal Lung III; Soares\_NhHMPu\_S1; Human Adult Pulmonary, re-excision; Osteoblasts; Human Cerebellum and Soares infant brain 1NIB.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.  
20 68 as residues: Pro-19 to Pro-26, Gly-33 to Ser-38, Gly-45 to Tyr-52, Ser-65 to Ser-71.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:18 and may have been publicly available prior to conception of  
25 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence

would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1084 of SEQ ID NO:18, b is an integer of 15 to 1098, where both a and b correspond to the positions of  
5 nucleotide residues shown in SEQ ID NO:18, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 9**

It has been discovered that this gene is expressed primarily in the following  
10 tissues/cDNA libraries: Prostate BPH; Soares\_pregnant\_uterus\_NbHPU and to a lesser extent in Jurkat Cells; H. Epididymus, caput & corpus; Activated T-cells; Human Infant Brain; Human Whole Six Week Old Embryo; Stratagene neuroepithelium (#937231); Resting T-Cell Library,II; Human Synovial Sarcoma; Bone marrow and Soares infant brain 1NIB.

15 Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 69 as residues: Pro-39 to Lys-54.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:19 and may have been publicly available prior to conception of  
20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1040 of SEQ ID NO:19, b  
25 is an integer of 15 to 1054, where both a and b correspond to the positions of

nucleotide residues shown in SEQ ID NO:19, and where b is greater than or equal to a + 14.

FEATURES OF PROTEIN ENCODED BY GENE NO: 10

5           The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil1255172 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "maternal transcript  
10   Maid [Mus musculus]". A partial alignment demonstrating the observed homology is shown immediately below.

```
>gi|1255172 maternal transcript Maid [Mus musculus] >sp|Q61162|Q61162
MATERNAL
15           INHIBITATION OF DIFFERENTIATION (MATERNAL TRANSCRIPT) (MEM4).
            Length = 310

            Plus Strand HSPs:

20           Score = 1134 (399.2 bits), Expect = 2.3e-114, P = 2.3e-114
            Identities = 232/310 (74%), Positives = 252/310 (81%), Frame = +2

            Q:    347 MFWRRLNEAAVTVSREATTLTIVFSQLPLPSPQETQKFCEQVHAAIKAFIAVYYLLPKDQ 526
                  MFWRRLNEAAV V+ EAT LT   FS+LP PSPQETQ+ CEQV   AI+   I VYY LPKDQ
25           S:    1 MFWRRLNEAAVKVNGEATVLTTHFSKLPWPSPQETQRICEQVRIAIEEIIIVYYSLPKDQ 60

            Q:    527 GITLRKLVRGATLDIVDGMAQLMEVLSVTPTQSPENNDLISYNSVWVACQQMPQIPRDNK 706
                  GITLRKLV R A LDIVDGMAQL+EVL   P+QS EN DLIS NSV VACQQ+P+IP+DNK
30           S:    61 GITLRKLV RNAALDIVDGMAQLLEVLLTAPSQSTENGDLISCNSVSVACQQVPEIPKDNK 120

            Q:    707 AAALLMLTKNVDFVKDAHEEMEQAVEECDPYSGLLXXXXXXXXXXXXXXXXXXVLGFPSNQD 886
                  AAALLMLTK+VDFVKDAHEEMEQAVEECDPYSGLL                   VLG PSN+D
            S:    121 AAALLMLTKSVDFVKDAHEEMEQAVEECDPYSGLLNDSEDNSDSHSEDEG-VLGLPSNRD 179

35           Q:    887 LYWSEDDQELIIPCLALVRASKACLKIRMLVAENGKKDQVAQLDDIVD-ISDEISPSVD 1063
                  YWSE+DQELIIPCLALVRAS+A LKKIR+LVAENGKKD+VAQLDDIV           +P
            S:    180 SYWSEEDQELIIPCLALVRASRASLKKIRILVAENGKKDEVAQLDDIVGHFLMRSAPVWI 239

40           Q:    1064 DLALSIYPPMCHLTVRINSAKLVSVLKKALEITKASHVTPQPEDSWIPLLINAIDHCMNR 1243
                  DL LSIYPP+CHLTVRI+SAKLVSVL KALEITKASHV+P P DSWIPLLINA+DHCMNR
            S:    240 DLVLSIYPPVCHLTVRISSAKLVSVLIKALEITKASHVSPHPGDSWIPLLINAVDHCMNR 299

            Q:    1244 IKELTQSELEL 1276
```

IK LTQ EL  
S: 300 IKALTQRAAEL 310

The segment of gil1255172 that is shown as "S" above is set out in the  
5 sequence listing as SEQ ID NO. 123. Based on the structural similarity these  
homologous polypeptides are expected to share at least some biological activities.  
Such activities are known in the art, some of which are described elsewhere herein.  
Assays for determining such activities are also known in the art, some of which have  
been described elsewhere herein.

10 Preferred polypeptides of the invention comprise a polypeptide having the  
amino acid sequence set out in the sequence listing as SEQ ID NO. 124 which  
corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following  
15 tissues/cDNA libraries: Soares fetal liver spleen 1NFLS and to a lesser extent in  
Soares melanocyte 2NbHM; Colon Normal III; Human 8 Week Whole Embryo;  
Human Osteoclastoma, re-excision; Human Thymus Stromal Cells; Hodgkin's  
Lymphoma II; Soares\_total\_fetus\_Nb2HF8\_9w; Morton Fetal Cochlea; Human  
Prostate Cancer, Stage C fraction; Synovial hypoxia-RSF subtracted; Myeloid  
20 Progenitor Cell Line; Fetal Liver, subtraction II; Human Prostate; Macrophage-  
oxLDL; NCI\_CGAP\_GCB1; NCI\_CGAP\_Kid3; HUMAN B CELL LYMPHOMA;  
Bone Marrow Cell Line (RS4,11); Normal lung; Smooth muscle control 2; Tongue  
carcinoma; Smooth muscle, control, re-excision; Supt Cells, cyclohexamide treated;  
Human Lung; Early Stage Human Lung, subtracted; Human Quadriceps; HSA 172  
25 Cells; Apoptotic T-cell, re-excision; B Cell lymphoma; Soares adult brain  
N2b4HB55Y; HEL cell line; pBMC stimulated w/ poly I/C; human corpus colosum;  
Human Osteoclastoma Stromal Cells - unamplified; Human Stomach, re-excision;

NCI\_CGAP\_GCB1; Synovial hypoxia; Synovial Fibroblasts (II1/TNF), subt; Human  
 Manic Depression Tissue; Human Neutrophil; T-Cell PHA 16 hrs; Apoptotic T-cell;  
 12 Week Old Early Stage Human, II; Human Pancreas Tumor; T-Cell PHA 24 hrs;  
 Human Heart; Liver, Hepatoma; Human umbilical vein endothelial cells, IL-4  
 5 induced; Synovial Fibroblasts (control); Hepatocellular Tumor, re-excision; Smooth  
 muscle, serum induced, re-exc; Macrophage-oxLDL, re-excision; Resting T-Cell  
 Library, II; Human T-Cell Lymphoma; Human Eosinophils; Colon Normal II; Human  
 Placenta; Adipocytes; H Macrophage (GM-CSF treated), re-excision; Colon Tumor  
 II; Human Placenta; NCI\_CGAP\_Co3; NCI\_CGAP\_Co9; NCI\_CGAP\_Pr6;  
 10 NCI\_CGAP\_GCB1; NCI\_CGAP\_Kid5; NCI\_CGAP\_PNS1; NCI\_CGAP\_Thy1;  
 Human Osteoclastoma; Human Bone Marrow, treated; Spleen, Chronic lymphocytic  
 leukemia; Activated T-cell(12h)/Thiouridine-re-excision; Human Testes;  
 Soares\_fetal\_lung\_NbHL19W; Stratagene hNT neuron (#937233); Keratinocyte;  
 Soares\_senescent\_fibroblasts\_NbHSF; Nine Week Old Early Stage Human;  
 15 Soares\_fetal\_lung\_NbHL19W and Soares infant brain 1NIB.

Many polynucleotide sequences, such as EST sequences, are publicly  
 available and accessible through sequence databases. Some of these sequences are  
 related to SEQ ID NO:20 and may have been publicly available prior to conception of  
 the present invention. Preferably, such related polynucleotides are specifically  
 20 excluded from the scope of the present invention. To list every related sequence  
 would be cumbersome. Accordingly, preferably excluded from the present invention  
 are one or more polynucleotides comprising a nucleotide sequence described by the  
 general formula of a-b, where a is any integer between 1 to 1653 of SEQ ID NO:20, b  
 is an integer of 15 to 1667, where both a and b correspond to the positions of  
 25 nucleotide residues shown in SEQ ID NO:20, and where b is greater than or equal to a  
 + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 11**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Primary Dendritic Cells, lib 1 and to a lesser extent in

5 Soares\_pregnant\_uterus\_NbHPU; Soares melanocyte 2NbHM; H.Leukocytes, normalized cot 50 B; Salivary Gland, Lib 3; Healing Abdomen wound, 70&90 min post incision; Smooth muscle, control, re-excision; HSA 172 Cells; Soares retina N2b4HR; Messangial cell, frac 2; Human Amygdala, re-excision; Synovial IL-1/TNF stimulated; NCI\_CGAP\_GCB1; H Female Bladder, Adult; Healing groin wound, 7.5

10 hours post incision; Synovial hypoxia-RSF subtracted; Human Colon, re-excision; H. Meningima, M1; Brain Frontal Cortex, re-excision; Human Adult Small Intestine; T-Cell PHA 16 hrs; Human Thymus Stromal Cells; Colon Tumor; Colon Carcinoma; Adipocytes; Colon Tumor II; NCI\_CGAP\_GCB1; Spleen, Chronic lymphocytic leukemia; Hodgkin's Lymphoma II and Soares placenta Nb2HP.

15 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:21 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence

20 would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1745 of SEQ ID NO:21, b is an integer of 15 to 1759, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:21, and where b is greater than or equal to a

25 + 14.

**FEATURES OF PROTEIN ENCODED BY GENE NO: 12**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares fetal liver spleen 1NFLS and to a lesser extent in Soares\_pregnant\_uterus\_NbHPU; Soares placenta Nb2HP; Activated T-Cell

5 (12hs)/Thiouridine labelledEco; Activated T-cell(12h)/Thiouridine-re-excision; HUMAN JURKAT MEMBRANE BOUND POLYSOMES; Human Adult Testes, Large Inserts, Reexcision; Soares\_fetal\_heart\_NbHH19W; Keratinocyte; Human Cerebellum; Amniotic Cells - Primary Culture; Stratagene ovarian cancer (#937219); Human endometrial stromal cells-treated with progesterone; Human Pituitary, subt

10 IX; Prostate BPH; H. Kidney Medulla, re-excision; Apoptotic T-cell; Human Pancreas Tumor; Human Rhabdomyosarcoma; Human Gall Bladder; Stratagene NT2 neuronal precursor 937230; Primary Dendritic cells,frac 2; Human Neutrophil, Activated; Endothelial-induced; Osteoblasts; Human Primary Breast Cancer; LNCAP + o.3nM R1881; Human Colon, subtraction; Human OB HOS treated (10 nM E2)

15 fraction I; Human Primary Breast Cancer,re-excision; Human Thyroid; Resting T-Cell, re-excision; Human Epididymus; Human Amygdala,re-excision; Human Colon Cancer,re-excision; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human Fetal Epithelium (Skin); Human Stomach,re-excision; Jurkat T-cell G1 phase; Human Manic Depression Tissue; Human Infant Brain; Breast Cancer Cell line, angiogenic;

20 Human Fetal Kidney; Soares\_fetal\_lung\_NbHL19W; Human Primary Breast Cancer Reexcision; Human Hypothalamus,Schizophrenia; Human Thymus; Bone Marrow Stromal Cell, untreated; Soares breast 2NbHBst; Human Adrenal Gland Tumor; Human adult testis, large inserts; NTERA2, control; Macrophage-oxLDL, re-excision; Adipocytes; Human Testes Tumor; Normal colon; Soares melanocyte

25 2NbHM; Human Testes, Reexcision; Human Placenta; Human Fetal Heart; human tonsils; Human Adult Pulmonary,re-excision; Human Amygdala; Human

Microvascular Endothelial Cells, fract. A; Monocyte activated; HUMAN B CELL LYMPHOMA; Spleen, Chronic lymphocytic leukemia; Human Testes; Human Endometrial Tumor; Hodgkin's Lymphoma II and Primary Dendritic Cells, lib 1.

Many polynucleotide sequences, such as EST sequences, are publicly  
5 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:22 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention  
10 are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 644 of SEQ ID NO:22, b is an integer of 15 to 658, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:22, and where b is greater than or equal to a + 14.

15

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 13**

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no.  
20 gnl|PID|e1348613 (all information available through the recited accession number is incorporated herein by reference) which is described therein as " Glutathione S-transferases protein [Caenorhabditis elegans]". A partial alignment demonstrating the observed homology is shown immediately below.

25 >gnl|PID|e1348613 similar to Glutathione S-transferases. [Caenorhabditis elegans] >sp|Q21925|Q21925 R11A8.5 PROTEIN.  
Length = 347

Plus Strand HSPs:



Score = 611 (215.1 bits), Expect = 6.1e-59, P = 6.1e-59  
 Identities = 124/267 (46%), Positives = 181/267 (67%), Frame = +3

```

5      Q:   153 LYQYKTCPFCSKVRAFLDFHALPYQVVEVNPVRRAEIKFSS-YRKVPILVAQEGESSQQL 329
          LYQY+TCPFC KVRAFLD+H   Y+VVEVNPV R++IKFS+ Y+KVPIL  + GE++  +
      S:   68 LYQYETCPFCKVRAFLDYHGFSYEVVEVNPVTRSQIKFSTTYKKVPIL--RSGETT--M 123

10     Q:   330 NDSSVIISALKTYLVS-GQPLEEIIITYYPAMKAVNEQGKEVTEFGNKYWLMLNEKEAQQV 506
          +S++IIS L TYL   Q L++II YPA+ + NE+GK V  + NK+++M  + +
      S:   124 TESTLIISTLATYLRPDQSLDQIIQMPAVDSTNEKGKPVNLNYPNKFFVMKGKVDGDAN 183

15     Q:   507 YGGKEARTEEMKWRQWADDWLVLHISPNNVYRTPTEALASFDYIVREGK----FGAVEGAV 674
          AR EE +WR+W D+W +HLISPNVYR   E++ +F +  + G   F A E  +
      S:   184 MAS--AR-EEREWEVDNWFHILISPNVYRNWNESVETFRWFQVGDWHRTFPAWERV 240

20     Q:   675 AKYMGAAAMYLIKRLKSRHRLQDNVREDLYEAADKWVAAVGKDRPFMGGQKPNLADLAV 854
          A Y+GAAAM+L+SK LK +H + D  RE+L +A   W+AA+G +R F+GG +PNLADL++
      S:   241 AVYVGAAAMFLLSKTLKKKHNNINDE-REELRKACRDWMAAIGPNRQFLGGDEPNLADLSL 299

      Q:   855 YGVLRVMEGLDAFDDLMQHTHIQPWYLRXERAI 953
          YG +   G  AF +++   I  W+ + + +
      S:   300 YGAMNSFYGCSAFKEVILEEKIAEWWRKMDALV 332
  
```

25       The segment of gnl|PID|e1348613 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 125. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have  
 30       been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 126 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

35       It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Macrophage-oxLDL; Human Cerebellum and to a lesser extent in Human Pituitary, subt IX; Human adult testis, large inserts; Human Eosinophils; Colon Carcinoma; Stratagene colon (#937204); Stratagene HeLa cell s3

937216; Human Fetal Heart; Human Aortic Endothelium; Human Fetal Bone; Human  
 Primary Breast Cancer; H. cerebellum, Enzyme subtracted; Smooth muscle, IL1b  
 induced; Salivary Gland, Lib 2; Soares\_fetal\_lung\_NbHL19W; Human Infant Brain;  
 Human Fetal Dura Mater; Soares\_fetal\_lung\_NbHL19W; Human Hippocampus;  
 5 Bone Marrow Stromal Cell, untreated; Soares breast 2NbHBst; HM1; Human adult  
 (K.Okubo); NCI\_CGAP\_Co9; NCI\_CGAP\_GC5; NCI\_CGAP\_Ov8;  
 NCI\_CGAP\_CLL1; NCI\_CGAP\_Kid5; NCI\_CGAP\_Pr22; NCI\_CGAP\_Brm25;  
 Adipocytes; Soares melanocyte 2NbHM; Human Adult Pulmonary, re-excision;  
 Activated T-Cell (12hs)/Thiouridine labelled Eco; Endothelial cells-control; Colon  
 10 Normal III; Human Bone Marrow, treated; Spleen, Chronic lymphocytic leukemia;  
 Activated T-cell(12h)/Thiouridine-re-excision; neutrophils control; T cell helper II;  
 Nine Week Old Early Stage Human and Primary Dendritic Cells, lib 1.

Many polynucleotide sequences, such as EST sequences, are publicly  
 available and accessible through sequence databases. Some of these sequences are  
 15 related to SEQ ID NO:23 and may have been publicly available prior to conception of  
 the present invention. Preferably, such related polynucleotides are specifically  
 excluded from the scope of the present invention. To list every related sequence  
 would be cumbersome. Accordingly, preferably excluded from the present invention  
 are one or more polynucleotides comprising a nucleotide sequence described by the  
 20 general formula of a-b, where a is any integer between 1 to 1449 of SEQ ID NO:23, b  
 is an integer of 15 to 1463, where both a and b correspond to the positions of  
 nucleotide residues shown in SEQ ID NO:23, and where b is greater than or equal to a  
 + 14.

25 **FEATURES OF PROTEIN ENCODED BY GENE NO: 14**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human Pituitary, subt IX and to a lesser extent in Human Pituitary; Human Pituitary, subtracted.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 5 74 as residues: Pro-24 to Gln-44.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:24 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically 10 excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 461 of SEQ ID NO:24, b is an integer of 15 to 475, where both a and b correspond to the positions of 15 nucleotide residues shown in SEQ ID NO:24, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 15**

The computer algorithm BLASTX has been used to determine that the 20 translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gnllPIDle1187970 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "F45G2.8 [Caenorhabditis elegans]". A partial alignment demonstrating the observed homology 25 is shown immediately below.

>gnl|PID|e1187970 F45G2.8 [Caenorhabditis elegans]

Length = 136

Plus Strand HSPs:

5      Score = 195 (68.6 bits), Expect = 5.8e-15, P = 5.8e-15  
       Identities = 40/71 (56%), Positives = 56/71 (78%), Frame = +3

Q:    210 QEAQQILNV-SKLSPEEVQKNYEHLPKVNDSVGGSFYLSKVVRAKERLDEEL-KIQAQ 383  
       +E+ QILNV + L+ EEV+K+YEHLF +NDKS GG+ YLSKVV RAKER+DEE +I+ +

10    S:    65 EESLQILNVKTPLNREEVEKHVEHLFNINDKSKGGTLYLSKVVRAKERIDEEFGRIELK 124

Q:    384 EDREKGQMPHT 416  
       E+++K +    T

15    S:    125 EEKKKEENAKT 135

The segment of gnl|PIDle1187970 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 127. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein.

20    Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 128 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a

25    sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: NCI\_CGAP\_Co3; Soares\_pregnant\_uterus\_NbHPU; Soares fetal liver spleen 1NFLS and to a lesser extent in Soares\_fetal\_lung\_NbHL19W; HM1; Rectum tumour; Frontal lobe,dementia,re-excision; Human Pineal Gland; HL-

30    60, PMA 4H, re-excision; Human Pituitary, subt IX; Human Adipose; Soares\_NhHMPu\_S1; Ulcerative Colitis; Human Fetal Brain; Human Synovial Sarcoma; Smooth muscle,control and Primary Dendritic Cells, lib 1.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are

related to SEQ ID NO:25 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 506 of SEQ ID NO:25, b is an integer of 15 to 520, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:25, and where b is greater than or equal to a + 14.

10

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 16

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no.

15 gill1628608 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "ubiquitin/ribosomal fusion protein [Sus scrofa]". A partial alignment demonstrating the observed homology is shown immediately below.

```

20   >gi|1628608 ubiquitin/ribosomal fusion protein [Sus scrofa] >gi|37565
    ubiquitin-52 amino acid fusion protein [Homo sapiens]
    >gi|37569
    ubiquitin-52 amino acid fusion protein [Homo sapiens]
    >gi|37567
25   ubiquitin-52 amino acid fusion protein [Homo sapiens]
    >gi|3288887
    (AC005253) ubiquitin-52 amino acid fusion protein [Homo
    sapiens]
    >gi|1050758 ubiquitin/ribosomal protein L40 [Rattus
30   norvegicus]
    >gi|600538 ubiquitin [artificial sequence] >gi|4139066
    (AF081142)
    ubiquitin/ribosomal protein CEP52 fusion protein [Cricetulus
    sp.]
35   >pir|S34428|UQHUR ubiquitin / ribosomal protein CEP52 - human

```

5                   >pir|I65237|I65237 ubiquitin/ribosomal protein L40 - rat  
                   >sp|Q95260|Q95260 UBIQUITIN/RIBOSOMAL FUSION PROTEIN.  
                   >sp|G3288887|G3288887 UBIQUITIN-52 AMINO ACID FUSION PROTEIN.  
                   >sp|G4139066|G4139066 UBIQUITIN/RIBOSOMAL PROTEIN CEP52 FUSION  
 10                   PROTEIN. >pir|S11248|S11248 ubiquitin / ribosomal protein  
                   CEP52 -  
                   mouse (fragment) {SUB 1-94} >gi|825728 ubiquitin precursor  
                   [Homo  
                   sapiens] {SUB 40-128} >gi|63867 ubiquitin-fusion protein (69  
 15                   AA).  
                   [Gallus gallus] {SUB 60-128} >gi|973178 ubiquitin/60S  
                   ribosomal  
                   subunit protein [Rattus norvegicus] {SUB 77-128}  
                   Length = 128  
 20                   Minus Strand HSPs:  
                   Score = 480 (169.0 bits), Expect = 4.9e-45, P = 4.9e-45  
                   Identities = 99/128 (77%), Positives = 105/128 (82%), Frame = -2  
 25                   Q:   813 MWIFVKTLTGQTITLEVEPSDATENVKAKIQDR\*GILPDQQCLIFEGKQLKDGHTLLDYN 634  
                   M IFVKTLTG+TITLEVEPSD   ENVKAKIQD+ GI PDQQ LIF GKQL+DG TL DYN  
                   S:    1 MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQLIFAGKQLEDGRTLSDYN 60  
 30                   Q:   633 IQKESTLRLMLCL\*GGITEPSLCHLAQKYNCKMMCHKSCARLRPHAVNCCKK-CGHTNN 457  
                   IQKESTL L+L L GGI EPSL   LAQKYNC KM+C K   ARL P AVNC KK CGHTNN  
                   S:    61 IQKESTLHLVLRLRGGIIEPSLRQLAQKYNC DKMICRKCYARLHPRAVNCRRKKCGHTNN 120  
                   Q:    456 LRPKKKVE 433  
                   LRPKKKV+  
                   S:    121 LRPKKKVK 128

The segment of gil1628608 that is shown as "S" above is set out in the  
 sequence listing as SEQ ID NO. 129. Based on the structural similarity these  
 35 homologous polypeptides are expected to share at least some biological activities.  
 Such activities are known in the art, some of which are described elsewhere herein.  
 Assays for determining such activities are also known in the art, some of which have  
 been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
 40 amino acid sequence set out in the sequence listing as SEQ ID NO. 130 which  
 corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in Human umbilical vein endothelial cells, IL-4 induced.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 76 as residues: Arg-21 to Ser-27, Ile-36 to Asp-41.

5 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:26 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence  
10 would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1040 of SEQ ID NO:26, b is an integer of 15 to 1054, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:26, and where b is greater than or equal to a  
15 + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 17**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: HE6W; Synovial hypoxia-RSF subtracted; Human  
20 endometrial stromal cells-treated with progesterone; Human Brain, Striatum; Human Umbilical Vein Endothelial Cells, uninduced; T-Cell PHA 24 hrs; Liver, Hepatoma; Epithelial-TNF $\alpha$  and INF induced; T Cell helper I; Osteoblasts; Primary Dendritic Cells, lib 1.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.  
25 77 as residues: Thr-40 to Asn-63.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:27 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 730 of SEQ ID NO:27, b is an integer of 15 to 744, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:27, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 18**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Colon Tumor II and to a lesser extent in Soares\_pregnant\_uterus\_NbHPU; Soares\_fetal\_liver\_spleen\_1NFLS\_S1; Soares\_multiple\_sclerosis\_2NbHMSP; Soares infant brain 1NIB; Soares ovary tumor NbHOT; Soares\_parathyroid\_tumor\_NbHPA; Colon Normal II; Soares retina N2b4HR; Raji Cells, cyclohexamide treated; normalized infant brain cDNA; Human Chondrosarcoma; Human Fetal Brain; Normal colon; Human Cerebellum; Soares fetal liver spleen 1NFLS; Infant brain, LLNL array of Dr. M. Soares 1NIB; Adipose tissue (human); Colon, normal; Dendritic Cells From CD34 Cells; Bone marrow stroma,treated; CD34+ cell, I, frac II; NCI\_CGAP\_Co3; NCI\_CGAP\_GC4; Ku 812F Basophils Line; Human Tonsil, Lib 3; Frontal Lobe, Dementia; Human Lung Cancer,re-excision; Human Epididymus; Breast Cancer cell line, MDA 36; Jurkat T-Cell, S phase; Human Ovary; Spinal Cord, re-excision; Human Fetal Dura Mater;



Human Uterine Cancer; Human fetal lung; Human colorectal cancer;  
Soares\_NhHMPu\_S1; Soares\_NFL\_T\_GBC\_S1; Soares\_total\_fetus\_Nb2HF8\_9w;  
Human Pancreas Tumor, Reexcision; 12 Week Old Early Stage Human; Brain frontal  
cortex; NCI\_CGAP\_Brn23; Pancreatic Islet; Soares melanocyte 2NbHM;  
5 Soares\_NhHMPu\_S1; Human Adult Pulmonary, re-excision; Spleen, Chronic  
lymphocytic leukemia and Keratinocyte.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.  
78 as residues: Ser-33 to Cys-43.

Many polynucleotide sequences, such as EST sequences, are publicly  
10 available and accessible through sequence databases. Some of these sequences are  
related to SEQ ID NO:28 and may have been publicly available prior to conception of  
the present invention. Preferably, such related polynucleotides are specifically  
excluded from the scope of the present invention. To list every related sequence  
would be cumbersome. Accordingly, preferably excluded from the present invention  
15 are one or more polynucleotides comprising a nucleotide sequence described by the  
general formula of a-b, where a is any integer between 1 to 1449 of SEQ ID NO:28, b  
is an integer of 15 to 1463, where both a and b correspond to the positions of  
nucleotide residues shown in SEQ ID NO:28, and where b is greater than or equal to a  
+ 14.

20

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 19**

It has been discovered that this gene is expressed primarily in the following  
tissues/cDNA libraries: H Female Bladder, Adult; Human Neutrophil; Anergic T-cell.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.  
25 79 as residues: Lys-23 to Lys-36.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:29 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 200 of SEQ ID NO:29, b is an integer of 15 to 214, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:29, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 20**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares melanocyte 2NbHM and to a lesser extent in Soares fetal liver spleen 1NFLS; Soares infant brain 1NIB; Soares retina N2b4HR; Human adult small intestine, re-excision; Human Whole Brain #2 - Oligo dT > 1.5Kb; Brain Frontal Cortex, re-excision; Apoptotic T-cell; Human Osteoblasts II; Human Hypothalamus, Schizophrenia; Soares adult brain N2b5HB55Y; Smooth muscle, serum treated and Soares\_fetal\_liver\_spleen\_1NFLS\_S1.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:30 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1145 of SEQ ID NO:30, b is an integer of 15 to 1159, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:30, and where b is greater than or equal to a + 14.

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 21**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares fetal liver spleen 1NFLS and to a lesser extent in

10 Soares infant brain 1NIB; Soares melanocyte 2NbHM; Human 8 Week Whole Embryo; Human Adult Retina; Nine Week Old Early Stage Human; Human Thymus; Soares adult brain N2b5HB55Y; Human Testes Tumor; Soares\_pregnant\_uterus\_NbHPU; Larynx carcinoma II; Morton Fetal Cochlea; Soares retina N2b4HR; Soares\_fetal\_lung\_NbHL19W;

15 Soares\_fetal\_heart\_NbHH19W; H Female Bladder, Adult; 12 Week Old Early Stage Human, II; T-Cell PHA 24 hrs; PERM TF274; Soares\_pregnant\_uterus\_NbHPU; Endothelial-induced; Human Amygdala; Soares\_fetal\_lung\_NbHL19W; Spleen, Chronic lymphocytic leukemia; Colorectal Tumor; LNCAP untreated; HUMAN TONSILS, FRACTION 2; Human Placenta; HUMAN STOMACH; Human T-cell

20 lymphoma,re-excision; Human adult small intestine,re-excision; Human Pineal Gland; B Cell lymphoma; Human Amygdala,re-excision; Human Colon Cancer,re-excision; Hepatocellular Tumor; Soares\_senescent\_fibroblasts\_NbHSF; Synovial IL-1/TNF stimulated; NTERA2 + retinoic acid, 14 days; Glioblastoma; Synovial Fibroblasts (IL1/TNF), subt; Brain Frontal Cortex, re-excision; Spinal Cord, re-

25 excision; Human Uterine Cancer; Human Hippocampus; Epithelial-TNF $\alpha$  and INF induced; Smooth muscle, serum induced,re-exc; Pancreas Islet Cell Tumor; Human

T-Cell Lymphoma; Colon Carcinoma; Human Synovial Sarcoma; Human Neutrophil, Activated; Activated T-Cell (12hs)/Thiouridine labelledEco; Soares\_fetal\_heart\_NbHH19W; Anergic T-cell; CD34 positive cells (Cord Blood); Stratagene HeLa cell s3 937216; Activated T-cell(12h)/Thiouridine-re-excision; Bone  
 5 Marrow Cell Line (RS4,11) and Human Cerebellum.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:31 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
 10 excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3102 of SEQ ID NO:31, b is an integer of 15 to 3116, where both a and b correspond to the positions of  
 15 nucleotide residues shown in SEQ ID NO:31, and where b is greater than or equal to a + 14.

## **FEATURES OF PROTEIN ENCODED BY GENE NO: 22**

It has been discovered that this gene is expressed primarily in the following  
 20 tissues/cDNA libraries: Soares fetal liver spleen 1NFLS and to a lesser extent in Soares retina N2b4HR; Soares\_fetal\_heart\_NbHH19W; CD34 positive cells (Cord Blood); HUMAN B CELL LYMPHOMA; Soares infant brain 1NIB; A1-CELL LINE; Soares retina N2b5HR; Healing groin wound, 6.5 hours post incision; Soares\_pregnant\_uterus\_NbHPU; Fetal Liver, subtraction II; Human Eosinophils;  
 25 Colon Tumor II; Bone marrow; CD34 depleted Buffy Coat (Cord Blood), re-excision; Soares\_fetal\_liver\_spleen\_1NFLS\_S1; T cell helper II; Human Pediatric Kidney;

CD34+ cell, I; Human Fetal Spleen; Resting T-Cell, re-excision;

Soares\_NhHMPu\_S1; Apoptotic T-cell, re-excision; Healing groin wound - zero hr post-incision (control); NTERA2 + retinoic acid, 14 days; Myeloid Progenitor Cell Line; Human Thymus; Human Rhabdomyosarcoma; Resting T-Cell Library,II;

- 5 Soares breast 3NbHBst; Soares\_fetal\_lung\_NbHL19W; Human Fetal Heart; Spleen, Chronic lymphocytic leukemia; Soares\_NhHMPu\_S1; Human 8 Week Whole Embryo; Human Cerebellum and Soares\_fetal\_lung\_NbHL19W.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:32 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 908 of SEQ ID NO:32, b is an integer of 15 to 922, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:32, and where b is greater than or equal to a + 14.

## 20 **FEATURES OF PROTEIN ENCODED BY GENE NO: 23**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human adult testis, large inserts; Human Testes; Soares placenta Nb2HP; Soares fetal liver spleen 1NFLS and to a lesser extent in Human Testes, Reexcision; LNCAP prostate cell line; Human 8 Week Whole Embryo; Soares infant brain 1NIB; Frontal Lobe, Dementia; Frontal lobe,dementia,re-excision; Smooth Muscle- HASTE normalized; T-Cell PHA 16 hrs; Human Adult Testes,

Large Inserts, Reexcision; Early Stage Human Brain; Human Fetal Heart and Human Microvascular Endothelial Cells, fract. A.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 83 as residues: Met-1 to Ser-6, Pro-23 to His-28, His-35 to Gln-40, Gln-52 to Pro-59,  
5 Pro-63 to Ser-69.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:33 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically  
10 excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2617 of SEQ ID NO:33, b is an integer of 15 to 2631, where both a and b correspond to the positions of  
15 nucleotide residues shown in SEQ ID NO:33, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 24**

The computer algorithm BLASTX has been used to determine that the  
20 translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil13011 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "URF 3 (NADH dehydrogenase subunit) [Homo sapiens]". A partial alignment demonstrating the  
25 observed homology is shown immediately below.

```
>gi|13011 URF 3 (NADH dehydrogenase subunit) [Homo sapiens] >gi|506832  
protein
```

3 [Homo sapiens] >pir|A00422|DNHUN3 NADH dehydrogenase  
 (ubiquinone)  
 (EC 1.6.5.3) chain 3 - human mitochondrion (SGC1)  
 >sp|P03897|NU3M\_HUMAN NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 3 (EC  
 5 1.6.5.3).  
 Length = 115

Minus Strand HSPs:

10 Score = 204 (71.8 bits), Expect = 5.1e-16, P = 5.1e-16  
 Identities = 45/91 (49%), Positives = 49/91 (53%), Frame = -3

Q: 320 PQLNGYIEKSTPYECGFDPI SPARVPFSIKFFLVAIXXXXXXXXXXXXXXXXXXXXXXXXX 141  
 PQLNGY+EKSTPYECGFDP+SPARVPFS+KFFLVAI

15 S: 25 PQLNGYMEKSTPYECGFDPM SPARVPFSMKFFLVAITFLLFDLEIALLLPLPWALQTTNL 84

Q: 140 XXIVMXXXXXXXXXXXXXXXXXXYE\*LQKGLD\*TE 48  
 +VM YE LQKGLD TE

S: 85 PLVMSSLLLIILALSLAYEWLQKGLDWTE 115

20

The segment of gil13011 that is shown as "S" above is set out in the sequence listing as SEQ ID NO: . Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

25

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 131 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

30

It has been discovered that this gene is expressed primarily in Osteoblasts.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:34 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence

35

would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 371 of SEQ ID NO:34, b is an integer of 15 to 385, where both a and b correspond to the positions of  
5 nucleotide residues shown in SEQ ID NO:34, and where b is greater than or equal to a + 14.

### **FEATURES OF PROTEIN ENCODED BY GENE NO: 25**

It has been discovered that this gene is expressed primarily in the following  
10 tissues/cDNA libraries: Soares infant brain 1NIB and to a lesser extent in Soares melanocyte 2NbHM; Soares fetal liver spleen 1NFLS; Stratagene colon (#937204); Soares\_pregnant\_uterus\_NbHPU; Soares\_pregnant\_uterus\_NbHPU; Stratagene colon (#937204); Human colon cancer, metatized to liver, subtraction; B Cell lymphoma; Human Amygdala, re-excision; H Female Bladder, Adult; Human Thymus; HUMAN  
15 JURKAT MEMBRANE BOUND POLYSOMES; Human Hypothalamus, Schizophrenia; Ulcerative Colitis; Epithelial-TNF $\alpha$  and INF induced; Human Testes Tumor; Endothelial cells-control and Human Amygdala.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
20 related to SEQ ID NO:35 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the  
25 general formula of a-b, where a is any integer between 1 to 2132 of SEQ ID NO:35, b is an integer of 15 to 2146, where both a and b correspond to the positions of



nucleotide residues shown in SEQ ID NO:35, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 26**

5 It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares retina N2b4HR and to a lesser extent in Ulcerative Colitis; Human Testes Tumor and Soares placenta Nb2HP.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
10 related to SEQ ID NO:36 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the  
15 general formula of a-b, where a is any integer between 1 to 1370 of SEQ ID NO:36, b is an integer of 15 to 1384, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:36, and where b is greater than or equal to a + 14.

#### **20 FEATURES OF PROTEIN ENCODED BY GENE NO: 27**

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil2564205 (all information available through the recited accession number is  
25 incorporated herein by reference, for example, J. Biol. Chem. 270 (24), 14801-14808 (1995)) which is described therein as "prenylated rab acceptor 1 [Rattus norvegicus]".

A partial alignment demonstrating the observed homology is shown immediately below.

```

5      >gi|2564205 (AF025506) prenylated rab acceptor 1 [Rattus norvegicus]
      Length = 185

      Plus Strand HSPs:

10     Score = 914 (321.7 bits), Expect = 4.1e-91, P = 4.1e-91
      Identities = 174/183 (95%), Positives = 178/183 (97%), Frame = +3

      Q:      3 AQKDQQKDAEAEGLSGTTLLPKLIPSGAGREWLERRRATIRPWSTFVDQQRFSRPRNLGE 182
      A QKDQQKDAE EGLS TTLLPKLIPSGAGREWLERRRATIRPW TFVDQQRFSRPRN+GE
      S:      3 AQKDQQKDAEVEGLSATLLPKLIPSGAGREWLERRRATIRPWGTFVDQQRFSRPRNVGE 62

15     Q:     183 LCQRLVRNVEYYQSNYVFVFLGLILYCVVTSPMLLVALAVFFGACYILYLRTLESKLVLF 362
      LCQRLVRNVEYYQSNYVFVFLGLILYCVVTSPMLLVALAVFFGACYILYLRTL+SKLVLF
      S:      63 LCQRLVRNVEYYQSNYVFVFLGLILYCVVTSPMLLVALAVFFGACYILYLRTLQSKLVLF 122

20     Q:     363 GREVSPA HQYALAGGISFPFFWLAGAGSAVFWVLGATLVVIGSHA AFHQIEAVDGEELQM 542
      GREVSPA HQYALAGG+SFPFFWLAGAGSAVFWVLGATLV+IGSHA AFHQIE DGEELQM
      S:     123 GREVSPA HQYALAGGV SFPFFWLAGAGSAVFWVLGATLV LIGSHA AFHQIEPADGEELQM 182

25     Q:     543 EPV 551
      EPV
      S:     183 EPV 185

```

The segment of gi2564205 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 132. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 133 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares melanocyte 2NbHM and to a lesser extent in Keratinocyte; Amniotic Cells - TNF induced; Human endometrial stromal cells-treated with progesterone; Soares adult brain N2b5HB55Y; HUMAN B CELL

5 LYMPHOMA; Stratagene pancreas (#937208); Human Primary Breast Cancer; Amniotic Cells - Primary Culture; Human Fetal Epithelium (Skin); Glioblastoma; Healing groin wound, 6.5 hours post incision;

Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human Pituitary, subt IX; Prostate BPH; Human Umbilical Vein Endothelial Cells, uninduced; Human Thymus; Macrophage

10 (GM-CSF treated); Human Whole Six Week Old Embryo; Human adult testis, large inserts; Stratagene liver (#937224); Early Stage Human Brain; Soares\_NhHMPu\_S1; Primary Dendritic cells,frac 2; Bone marrow; Human Fetal Heart; Endothelial cells-control; Bone Marrow Cell Line (RS4,11); Stratagene hNT neuron (#937233); Osteoblasts; Soares placenta Nb2HP; Testes; Human Kidney; Human Colon,

15 subtraction; H. Striatum Depression, subt; Human OB HOS control fraction I; stomach cancer (human); Human White Adipose; Human Primary Breast Cancer, re-excision; Human Pituitary, subtracted; Smooth Muscle Serum Treated, Norm; Human Fetal Bone; Smooth muscle-ILb induced; Aorta endothelial cells + TNF-a; H. Epididymus, cauda; Human Pineal Gland; Human Epididymus; Human endometrial

20 stromal cells-treated with estradiol; human corpus colosum; Synovial hypoxia-RSF subtracted; Human Osteosarcoma; Synovial hypoxia; Stratagene pancreas (#937208); Myeloid Progenitor Cell Line; Soares\_fetal\_heart\_NbHH19W; Soares\_senescent\_fibroblasts\_NbHSF; H. Meningioma, M1; H. Lymph node breast Cancer; Spinal Cord, re-excision; T-Cell PHA 16 hrs; Human Bone Marrow, re-

25 excision; Human Brain, Striatum; L428; Human Fetal Kidney; human ovarian cancer; Human Fetal Dura Mater; T-Cell PHA 24 hrs; Ulcerative Colitis; Human Adrenal

Gland Tumor; Pancreas Islet Cell Tumor; Human Substantia Nigra;  
Soares\_placenta\_8to9weeks\_2NbHP8to9W; NCI\_CGAP\_GCB1; Human Fetal Lung  
III; Endothelial-induced; Anergic T-cell; Smooth muscle, control; Spleen, Chronic  
lymphocytic leukemia; Human Endometrial Tumor; Hodgkin's Lymphoma II; T cell  
5 helper II; Nine Week Old Early Stage Human; Human Cerebellum and Primary  
Dendritic Cells, lib 1.

The gene encoding the disclosed cDNA is believed to reside on chromosome  
19. Accordingly, polynucleotides related to this invention are useful as a marker in  
linkage analysis for chromosome 19.

10 The tissue distribution in melanocytes suggests the protein product of this  
clone is useful for the treatment, diagnosis, and/or prevention of various skin  
disorders including congenital disorders (i.e. nevi, moles, freckles, Mongolian spots,  
hemangiomas, port-wine syndrome), integumentary tumors (i.e. keratoses, Bowen's  
disease, basal cell carcinoma, squamous cell carcinoma, malignant melanoma, Paget's  
15 disease, mycosis fungoides, and Kaposi's sarcoma), injuries and inflammation of the  
skin (i.e. wounds, rashes, prickly heat disorder, psoriasis, dermatitis), atherosclerosis,  
urticaria, eczema, photosensitivity, autoimmune disorders (i.e. lupus erythematosus,  
vitiligo, dermatomyositis, morphea, scleroderma, pemphigoid, and pemphigus),  
keloids, striae, erythema, petechiae, purpura, and xanthelasma. In addition, such  
20 disorders may predispose increased susceptibility to viral and bacterial infections of  
the skin (i.e. cold sores, warts, chickenpox, molluscum contagiosum, herpes zoster,  
boils, cellulitis, erysipelas, impetigo, tinea, athlete's foot, and ringworm). Moreover,  
the protein product of this clone may also be useful for the treatment or diagnosis of  
various connective tissue disorders such as arthritis, trauma, tendonitis,  
25 chondromalacia and inflammation, autoimmune disorders such as rheumatoid  
arthritis, lupus, scleroderma, and dermatomyositis as well as dwarfism, spinal

deformation, and specific joint abnormalities as well as chondrodysplasias (i.e. spondyloepiphyseal dysplasia congenita, familial osteoarthritis, Atelosteogenesis type II, metaphyseal chondrodysplasia type Schmid). This protein is believed to be a novel member of the Rab acceptor protein family. Rab proteins are members of the Ras super gene family. Mutations in the Ras gene have been correlated to transformation of cells and tissues to cancer phenotypes and genotypes. Rab, specifically, is a GTP binding protein which plays a role in intra-Golgi transport. Thus, this protein may be useful as an agonist or antagonist of signal transduction pathways, particularly those involved in Golgi transport, and the protein may play either a direct or indirect role. Therefore, this protein is useful in the detection, treatment, and/or prevention of developmental, proliferative, and cancerous diseases and/or disorders. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:37 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 725 of SEQ ID NO:37, b is an integer of 15 to 739, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:37, and where b is greater than or equal to a + 14.

25

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 28**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human Ovary and to a lesser extent in Healing Abdomen wound, 70&90 min post incision; Human Colon Cancer, re-excision; Human Adipose Tissue, re-excision; Soares\_pregnant\_uterus\_NbHPU; Ulcerative Colitis; Soares breast 2NbHBst; 12 Week Old Early Stage Human; Nine Week Old Early Stage Human and Primary Dendritic Cells, lib 1.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 88 as residues: Cys-13 to Trp-19, Ser-22 to Ile-27.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:38 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 729 of SEQ ID NO:38, b is an integer of 15 to 743, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:38, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 29**

It has been discovered that this gene is expressed primarily in Resting T-Cell Library, II.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:39 and may have been publicly available prior to conception of

the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the  
5 general formula of a-b, where a is any integer between 1 to 1052 of SEQ ID NO:39, b is an integer of 15 to 1066, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:39, and where b is greater than or equal to a + 14.

#### 10 **FEATURES OF PROTEIN ENCODED BY GENE NO: 30**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares infant brain 1N1B and to a lesser extent in Early Stage Human Brain; Soares\_NhHMPu\_S1; Duodenum; Frontal Lobe, Dementia; Soares\_NhHMPu\_S1; H. Striatum Depression, subt; Frontal lobe,dementia,re-  
15 excision; Spinal Cord, re-excision; Stromal cell TF274; Spinal cord; Human Fetal Brain and Soares\_pineal\_gland\_N3HPG.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 90 as residues: Tyr-59 to Arg-64, Pro-101 to Met-116.

Many polynucleotide sequences, such as EST sequences, are publicly  
20 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:40 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention  
25 are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1130 of SEQ ID NO:40, b

is an integer of 15 to 1144, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:40, and where b is greater than or equal to a + 14.

## **5 FEATURES OF PROTEIN ENCODED BY GENE NO: 31**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Human adult small intestine, re-excision; Human Adult Small Intestine; Human Whole Six Week Old Embryo; Soares melanocyte 2NbHM; Human Fetal Lung III; Activated T-cell(12h)/Thiouridine-re-excision.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:41 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence  
15 would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1173 of SEQ ID NO:41, b is an integer of 15 to 1187, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:41, and where b is greater than or equal to a  
20 + 14.

## **FEATURES OF PROTEIN ENCODED BY GENE NO: 32**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: PERM TF274.

25 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are



related to SEQ ID NO:42 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1206 of SEQ ID NO:42, b is an integer of 15 to 1220, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:42, and where b is greater than or equal to a + 14.

10

### FEATURES OF PROTEIN ENCODED BY GENE NO: 33

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no.

15 gil2197085 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "ORF2-like protein [Homo sapiens]". A partial alignment demonstrating the observed homology is shown immediately below.

```

20  >gi|2197085 (AF003535) ORF2-like protein [Homo sapiens]
      Length = 573

      Minus Strand HSPs:

25  Score = 311 (109.5 bits), Expect = 8.2e-40, Sum P(2) = 8.2e-40
      Identities = 98/246 (39%), Positives = 131/246 (53%), Frame = -1

      Q: 1285 PIPLKLFQKTEKKLC-NTFYEASITWIPKPDKYTTRKDK-----LDTKIFYLXKY* 1133
          P LKLFQ E++ + N+FYEASI IPKP + TT+K+ +D KI L K
30  S: 4 PFLKLFQSIKEGILPNSFYEASIIIPKPGRDITTKKENFRPISLMNIDAKI--LNKIL 61

      Q: 1132 *AKFSSILKRIIYHDHMEFISRIQGWFSXRV SINVIHHINTMKDKKSM-----QSFQF 974
          + +K++I+HD + FI +QGWF+ R SINVI HIN KDK M ++F
35  S: 62 ANRIQQHIKKLIHHDQVGFIPGMQGWFNIRKSINVIQHINRAKDKNHMIISIDA EKAFDK 121

      Q: 973 TQNKHLTKFYNL SG*KHGTFKMC*RGFVKNPQL-IYLMRKKTEKFLP*QRVRYGCMLYPV 797

```

Q + K N G FK+ R P I L +K E F R GC L P+  
 S: 122 IQQPFMLKTLNKLGDGMFKII-RAIYDKPTANIILNGQKLEAFPLKTGTRQGCPLSPL 180  
 Q: 796 LLNNVLKVLTMIRQINKQTGK\*IASKLERKQ\*NSLFNND-IFYLENQRI--HKLVELIN 626  
 5 L N VL+VL IRQ + G + K E K SLF +D I YLEN + L++LI+  
 S: 181 LFNIVLEVLARAIRQEKEIKGIQLG-KEEVKL--SLFADDMIVYLENPIVSAQNLLKLIS 237  
 Q: 625 KSSKIAGYMFNI 590  
 SK++GY N+  
 10 S: 238 NFSKVSGYKINV 249

The segment of gil2197085 that is shown as "S" above is set out in the  
 sequence listing as SEQ ID NO. 134. Based on the structural similarity these  
 homologous polypeptides are expected to share at least some biological activities.  
 15 Such activities are known in the art, some of which are described elsewhere herein.  
 Assays for determining such activities are also known in the art, some of which have  
 been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
 amino acid sequence set out in the sequence listing as SEQ ID NO. 135 which  
 20 corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following  
 tissues/cDNA libraries: PERM TF274.

Many polynucleotide sequences, such as EST sequences, are publicly  
 25 available and accessible through sequence databases. Some of these sequences are  
 related to SEQ ID NO:43 and may have been publicly available prior to conception of  
 the present invention. Preferably, such related polynucleotides are specifically  
 excluded from the scope of the present invention. To list every related sequence  
 would be cumbersome. Accordingly, preferably excluded from the present invention  
 30 are one or more polynucleotides comprising a nucleotide sequence described by the  
 general formula of a-b, where a is any integer between 1 to 1271 of SEQ ID NO:43, b

is an integer of 15 to 1285, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:43, and where b is greater than or equal to a + 14.

## 5 FEATURES OF PROTEIN ENCODED BY GENE NO: 34

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: PERM TF274 and to a lesser extent in HL-60, PMA 4H, re-excision.

Many polynucleotide sequences, such as EST sequences, are publicly  
10 available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:44 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention:  
15 are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 971 of SEQ ID NO:44, b is an integer of 15 to 985, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:44, and where b is greater than or equal to a + 14.

20

## FEATURES OF PROTEIN ENCODED BY GENE NO: 35

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Hemangiopericytoma and to a lesser extent in breast lymph node CDNA library; Activated T-cell(12h)/Thiouridine-re-excision;  
25 NCI\_CGAP\_Pr22; Human Synovial Sarcoma; Human Bone Marrow, treated; Rejected Kidney, lib 4; HUMAN B CELL LYMPHOMA; Human Cerebellum;

Primary Dendritic Cells, lib 1; Soares ovary tumor NbHOT; NCI\_CGAP\_Co3;  
 Human Primary Breast Cancer; Healing groin wound - zero hr post-incision (control);  
 L428; Human Uterine Cancer; Human Adult Testes, Large Inserts, Reexcision; T-Cell  
 PHA 24 hrs; Soares\_fetal\_heart\_NbHH19W; Human adult testis, large inserts; Fetal  
 5 Heart; Human Fetal Kidney, Reexcision; Stratagene colon (#937204); Primary  
 Dendritic cells,frac 2; CD34 depleted Buffy Coat (Cord Blood), re-excision; Colon  
 Normal III; Bone Marrow Cell Line (RS4,11); Human Colon Cancer; Human Colon  
 Cancer, differential; H. hypothalamus, frac A, re-excision; Gessler Wilms tumor;  
 HM3; Soares\_testis\_NHT; NCI\_CGAP\_Br3; NCI\_CGAP\_GC5; Lung Carcinoma  
 10 A549 TNFalpha activated; Ku 812F Basophils Line; Human Prostate Cancer, Stage  
 B2 fraction; LNCAP untreated; H. Striatum Depression, subt; Human Primary Breast  
 Cancer, re-excision; Human Adult Retina; Human Pituitary, subtracted; Frontal  
 lobe, dementia, re-excision; Supt Cells, cyclohexamide treated; H. Epididymus, caput  
 & corpus; Aorta endothelial cells + TNF-a; Human Soleus; NTERA2 teratocarcinoma  
 15 cell line+retinoic acid (14 days); Human Pineal Gland; Human Tonsils, Lib 2;  
 STROMAL -OSTEOCLASTOMA; Soares\_fetal\_lung\_NbHL19W; Synovial IL-  
 1/TNF stimulated; Breast Cancer cell line, MDA 36; human corpus colosum; Healing  
 groin wound, 7.5 hours post incision; Amniotic Cells - Primary Culture; Salivary  
 Gland, Lib 2; Healing groin wound, 6.5 hours post incision; HL-60, PMA 4H, re-  
 20 excision; Human Pituitary, subt IX; wilm's tumor; H. Lymph node breast Cancer;  
 Breast Cancer Cell line, angiogenic; NCI\_CGAP\_Kid3; Monocyte activated, re-  
 excision; Human Umbilical Vein Endothelial Cells, uninduced; Human Primary  
 Breast Cancer Reexcision; Human Thymus Stromal Cells; Human Thymus; Human  
 Whole Six Week Old Embryo; NCI\_CGAP\_Kid6; Resting T-Cell Library, II; H.  
 25 Frontal cortex, epileptic, re-excision; Human T-Cell Lymphoma; Colon Carcinoma;  
 Smooth muscle, serum treated; Colon Normal II; Human Placenta; Colon Tumor II;

Soares melanocyte 2NbHM; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human  
Fetal Heart; human tonsils; Endothelial cells-control; Smooth muscle,control; Spleen;  
Chronic lymphocytic leukemia; Soares\_fetal\_lung\_NbHL19W; Hodgkin's  
Lymphoma II; Stratagene pancreas (#937208); Osteoblasts; Stratagene pancreas  
5 (#937208) and Keratinocyte.

Many polynucleotide sequences, such as EST sequences, are publicly  
available and accessible through sequence databases. Some of these sequences are  
related to SEQ ID NO:45 and may have been publicly available prior to conception of  
the present invention. Preferably, such related polynucleotides are specifically  
10 excluded from the scope of the present invention. To list every related sequence  
would be cumbersome. Accordingly, preferably excluded from the present invention  
are one or more polynucleotides comprising a nucleotide sequence described by the  
general formula of a-b, where a is any integer between 1 to 2596 of SEQ ID NO:45, b  
is an integer of 15 to 2610, where both a and b correspond to the positions of  
15 nucleotide residues shown in SEQ ID NO:45, and where b is greater than or equal to a  
+ 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 36**

It has been discovered that this gene is expressed primarily in the following  
20 tissues/cDNA libraries: PERM TF274.

Many polynucleotide sequences, such as EST sequences, are publicly  
available and accessible through sequence databases. Some of these sequences are  
related to SEQ ID NO:46 and may have been publicly available prior to conception of  
the present invention. Preferably, such related polynucleotides are specifically  
25 excluded from the scope of the present invention. To list every related sequence  
would be cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 937 of SEQ ID NO:46, b is an integer of 15 to 951, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:46, and where b is greater than or equal to a + 14.

### FEATURES OF PROTEIN ENCODED BY GENE NO: 37

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil13011 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "URF 3 (NADH dehydrogenase subunit) [Homo sapiens]". A partial alignment demonstrating the observed homology is shown immediately below.

```

>gi|13011 URF 3 (NADH dehydrogenase subunit) [Homo sapiens] >gi|506832
protein
      3 [Homo sapiens] >pir|A00422|DNHUN3 NADH dehydrogenase
(ubiquinone)
      (EC 1.6.5.3) chain 3 - human mitochondrion (SGC1)
      >sp|P03897|NU3M_HUMAN NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 3 (EC
      1.6.5.3).
      Length = 115

Minus Strand HSPs:

Score = 212 (74.6 bits), Expect = 9.9e-17, P = 9.9e-17
Identities = 40/45 (88%), Positives = 44/45 (97%), Frame = -1

Q:   329 LLIIITF*LPQLNGYIEKSTPYECGFDPISPARVPFSIKFFLVAI 195
      LL+IITF LPQLNGY+EKSTPYECGFDP+SPARVPFS+KFFLVAI
S:   16  LLMIITFWLPQLNGYMEKSTPYECGFDPMSPARVPFSMKFFLVAI 60

```

The segment of gil13011 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 136. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities

are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
5 amino acid sequence set out in the sequence listing as SEQ ID NO. 137 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in H. Striatum Depression, subt.

10 Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 97 as residues: Lys-17 to Leu-23.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:47 and may have been publicly available prior to conception of  
15 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 540 of SEQ ID NO:47, b  
20 is an integer of 15 to 554, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:47, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 38**

25 The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting

example, the sequence accessible through the following database accession no.  
 gil3135473 (all information available through the recited accession number is  
 incorporated herein by reference, for example, Genomics 51 (1), 154-157 (1998))  
 which is described therein as "Cre binding protein-like 2 [Homo sapiens]". A partial  
 5 alignment demonstrating the observed homology is shown immediately below.

```

10      >gi|3135473 (AF039081) Cre binding protein-like 2 [Homo sapiens]
        >sp|O60519|O60519 CRE BINDING PROTEIN-LIKE 2.
        Length = 120

        Plus Strand HSPs:

        Score = 508 (178.8 bits), Expect = 3.2e-47, P = 3.2e-47
        Identities = 103/120 (85%), Positives = 103/120 (85%), Frame = +3
15
        Q:   318 MDDSXXXXXXXXXXXXXXXXXXXXAKIDLKAKLERSRQSARECRARKKLRYQYLEELVSSRER 497
              MDDS                                AKIDLKAKLERSRQSARECRARKKLRYQYLEELVSSRER
        S:    1 MDDSKVVGKVKPKPGKRGRKPAKIDLKAKLERSRQSARECRARKKLRYQYLEELVSSRER 60

20      Q:   498 AICALREELEMYKQWCMAMDQGKIPSEIKALLTGEEQNKSQQNSSRHTKAGKTDANSNSW 677
              AICALREELEMYKQWCMAMDQGKIPSEIKALLTGEEQNKSQQNSSRHTKAGKTDANSNSW
        S:    61 AICALREELEMYKQWCMAMDQGKIPSEIKALLTGEEQNKSQQNSSRHTKAGKTDANSNSW 120
  
```

The segment of gil3135473 that is shown as "S" above is set out in the  
 25 sequence listing as SEQ ID NO. 138. Based on the structural similarity these  
 homologous polypeptides are expected to share at least some biological activities.  
 Such activities are known in the art, some of which are described elsewhere herein.  
 Assays for determining such activities are also known in the art, some of which have  
 been described elsewhere herein.

30 Preferred polypeptides of the invention comprise a polypeptide having the  
 amino acid sequence set out in the sequence listing as SEQ ID NO. 139 which  
 corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following  
 35 tissues/cDNA libraries: Primary Dendritic Cells, lib 1 and to a lesser extent in Soares



infant brain 1NIB; Morton Fetal Cochlea; Human T-cell lymphoma, re-excision;  
 Soares\_multiple\_sclerosis\_2NbHMSP; Stratagene NT2 neuronal precursor 937230;  
 Soares melanocyte 2NbHM; Soares\_NhHMPu\_S1; Primary Dendritic cells, frac 2;  
 Human Amygdala; Frontal Lobe, Dementia; Soares\_NhHMPu\_S1; Spinal Cord, re-  
 5 excision; Temporal cortex-Alzheimer, subtracted; Stratagene colon (#937204);  
 Human Testes Tumor; Human Bone Marrow, treated; Human Endometrial Tumor;  
 Hodgkin's Lymphoma II; Soares\_fetal\_heart\_NbHH19W; Soares fetal liver spleen  
 1NFLS; H. Brain, X-Chromosome hybridization; Poly[I]/Poly[C] Normal Lung  
 Fibroblasts; Kidney Cortex; Colon, normal; Human osteoarthritic, fraction II; H.  
 10 hypothalamus, frac A; Human Hippocampus, subtracted; H. Adipose Tissue; Soares  
 retina N2b5HR; Smooth muscle-ILb induced; Human Thyroid; Human Lung; Human  
 Soleus; Human Pineal Gland; Morton Fetal; Hepatocellular Tumor; Human Prostate  
 Cancer, Stage C fraction; LNCAP prostate cell line; Ovarian Tumor 10-3-95;  
 Pancreas normal PCA4 No; Human Manic Depression Tissue; Stratagene hNT neuron  
 15 (#937233); Stratagene neuroepithelium (#937231); Human Prostate; T-Cell PHA 16  
 hrs; L428; Apoptotic T-cell; Human Hypothalamus, Schizophrenia; Liver, Hepatoma;  
 Stratagene NT2 neuronal precursor 937230; Human Adipose; Spinal cord; Bone  
 Marrow Stromal Cell, untreated; Rejected Kidney, lib 4; H. Frontal  
 cortex, epileptic, re-excision; Human T-Cell Lymphoma; Dendritic cells, pooled;  
 20 Normal colon; Human Fetal Lung III; human tonsils; Spleen, Chronic lymphocytic  
 leukemia; Human Testes; Soares\_fetal\_lung\_NbHL19W;  
 Soares\_senescent\_fibroblasts\_NbHSF; Human 8 Week Whole Embryo and Soares  
 placenta Nb2HP.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO.

25 98 as residues: Met-28 to Gln-34.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:48 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 3038 of SEQ ID NO:48, b is an integer of 15 to 3052, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:48, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 39

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gnl|PID|e1347752 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "guanine nucleotide binding proteins [Caenorhabditis elegans]". A partial alignment demonstrating the observed homology is shown immediately below.

```

>gnl|PID|e1347752 weak similarity to guanine nucleotide binding proteins
    [Caenorhabditis elegans] >sp|Q21115|Q21115 K02B12.3 PROTEIN.
    Length = 365

25      Plus Strand HSPs:

      Score = 127 (44.7 bits), Expect = 5.3e-09, Sum P(2) = 5.3e-09
      Identities = 31/103 (30%), Positives = 54/103 (52%), Frame = +2

30      Q:  1085 SCGHEVVSCLDVSESGTFLGLGTVTGSAIYIAFSLQCLYYVREAHGIVVTDVAFL---- 1252
           S G + +S L VS+ G F  +GT++GSV ++      + LY+  E+HG+ VT + F+

```

S: 238 SRGFKSISSLAVSDCGNFTAVGTMSGSVLVFDTHECRRLYFSPESHGLFVTGIEFVSRTS 297

Q: 1253 PE-----KGRGPELLGSHETALFSVAVDSRCQLHLLPSRRSVP 1366

5 S: 298 PSICEDIQSETPGIASGFQSAVVTLAADKTMQLHRVPYPQPQP 340

P + P + ++A+ ++A D QLH +P + P

The segment of gnl|PID|e1347752 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 140. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities.

10 Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO.

15 <SEQIDNO:PR043:39:2> which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares ovary tumor NbHOT and to a lesser extent in Soares infant brain 1NIB; Soares placenta Nb2HP; Primary Dendritic Cells, lib 1; Soares\_fetal\_lung\_NbHL19W; Human Pituitary, subt IX; Soares\_pregnant\_uterus\_NbHPU; Colon Carcinoma; NCI\_CGAP\_GCB1; Activated T-cell(12h)/Thiouridine-re-excision; Human Cerebellum; Soares\_pregnant\_uterus\_NbHPU; NCI\_CGAP\_GCB1; Dendritic cells, pooled; 25 Human Adult Pulmonary, re-excision; Hodgkin's Lymphoma II; Soares fetal liver spleen 1NFLS; Human Pituitary, subtracted V; Human Pituitary, subtracted VIII; Human kidney Cortex, subtracted; H.Leukocytes, normalized cot 50A3; Human colorectal cancer; Soares\_parathyroid\_tumor\_NbHPA; Soares\_fetal\_liver\_spleen\_1NFLS\_S1; Soares\_senescent\_fibroblasts\_NbHSF;

Stratagene lung carcinoma 937218; Stratagene cDNA library Human heart,  
 cat#936208; Colon Normal; Activated T-Cells, 8 hrs., ligation 2;  
 Soares\_fetal\_liver\_spleen\_1NFLS\_S1; HepG2 Cells, lambda library; Human OB  
 HOS treated (1 nM E2) fraction I; Soares retina N2b5HR; Human Fetal Brain;  
 5 Human Cerebellum, subtracted; Activated T-cells; Human Adult Heart, re-excision;  
 Alzheimers, spongy change; Healing groin wound, 6.5 hours post incision; Jurkat T-  
 cell G1 phase; wilm's tumor; Human Infant Brain; Human Umbilical Vein,  
 Reexcision; Breast Cancer Cell line, angiogenic; CHME Cell Line, treated 5 hrs;  
 L428; Apoptotic T-cell; human ovarian cancer; Human Uterine Cancer; Human  
 10 Hippocampus; Soares\_pregnant\_uterus\_NbHPU; Liver, Hepatoma; Stratagene colon  
 (#937204); Human Adipose; Soares\_fetal\_lung\_NbHL19W; Macrophage (GM-CSF  
 treated); Pancreas Islet Cell Tumor; Human Testes Tumor; H Macrophage (GM-CSF  
 treated), re-excision; Normal colon; Soares melanocyte 2NbHM; human tonsils;  
 Endothelial-induced; Human adult (K.Okubo); NCI\_CGAP\_Br2; NCI\_CGAP\_Co3;  
 15 NCI\_CGAP\_GC1; NCI\_CGAP\_GC2; NCI\_CGAP\_Lu5; NCI\_CGAP\_Co10;  
 NCI\_CGAP\_Co11; NCI\_CGAP\_Kid3; Colon Normal III; HUMAN B CELL  
 LYMPHOMA and Keratinocyte.

Many polynucleotide sequences, such as EST sequences, are publicly  
 available and accessible through sequence databases. Some of these sequences are  
 20 related to SEQ ID NO:49 and may have been publicly available prior to conception of  
 the present invention. Preferably, such related polynucleotides are specifically  
 excluded from the scope of the present invention. To list every related sequence  
 would be cumbersome. Accordingly, preferably excluded from the present invention  
 are one or more polynucleotides comprising a nucleotide sequence described by the  
 25 general formula of a-b, where a is any integer between 1 to 2140 of SEQ ID NO:49, b  
 is an integer of 15 to 2154, where both a and b correspond to the positions of

nucleotide residues shown in SEQ ID NO:49, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 40**

5           It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Activated T-cells, 24 hrs, re-excision and to a lesser extent in Activated T-Cells, 24 hrs.; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human epithelioid sarcoma; Human Pituitary, re-excision; Human Aortic Endothelium; Human Primary Breast Cancer, re-excision; Human Thyroid; NCI\_CGAP\_GCB1; 10   NCI\_CGAP\_Kid1; NCI\_CGAP\_Kid6; Human Tonsils, Lib 2; Healing groin wound, 6.5 hours post incision; Rejected Kidney, lib 4; Soares breast 3NbHBst; Human Synovial Sarcoma; Soares\_multiple\_sclerosis\_2NbHMSP and CD34 depleted Buffy Coat (Cord Blood), re-excision.

          Many polynucleotide sequences, such as EST sequences, are publicly 15   available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:50 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention 20   are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 629 of SEQ ID NO:50, b is an integer of 15 to 643, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:50, and where b is greater than or equal to a + 14.

25

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 41**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Activated T-cells, 24 hrs, re-excision and to a lesser extent in Activated T-Cells, 24 hrs.; Soares fetal liver spleen 1NFLS; Soares\_fetal\_heart\_NbHH19W; Synovial hypoxia-RSF subtracted; Human Chronic  
5 Synovitis; Stratagene fetal retina 937202; T-Cell PHA 24 hrs; Hemangiopericytoma; Spleen, Chronic lymphocytic leukemia; Human Testes and Primary Dendritic Cells, lib 1.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 101 as residues: Glu-1 to Ser-9, Ser-23 to Ser-35.

10 Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:51 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence  
15 would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 668 of SEQ ID NO:51, b is an integer of 15 to 682, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:51, and where b is greater than or equal to a  
20 + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 42**

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting  
25 example, the sequence accessible through the following database accession no. gil968973 (all information available through the recited accession number is

incorporated herein by reference, for example, Proc. Natl. Acad. Sci. U.S.A. 92 (21), 9530-9534 (1995)) which is described therein as "epidermal growth factor receptor [Mus musculus]". A partial alignment demonstrating the observed homology is shown immediately below.

```

5      >gi|968973 involved in signaling by the epidermal growth factor receptor;
      Method: conceptual translation supplied by author [Mus
musculus]
      >sp|Q60902|Q60902 EPIDERMAL GROWTH FACTOR PATHWAY SUBSTRATE 15,
10     RELATED SEQUENCE (EPS15R).
      Length = 907

      Plus Strand HSPs:

15     Score = 1594 (561.1 bits), Expect = 1.1e-249, Sum P(3) = 1.1e-249
      Identities = 319/372 (85%), Positives = 332/372 (89%), Frame = +3

      Q:   492 QEVKEIFMHSGLTQNLLAHIWALADTRQTGKLSKDQFALAMYFIQQKVS KGIDPPQVLSP 671
      QEVKEIFMHSGLTQNLLAHIWALADTRQTGKLSK+QFALAMYFIQQKVS KGIDPPQVLSP
20     S:   295 QEVKEIFMHSGLTQNLLAHIWALADTRQTGKLSKEQFALAMYFIQQKVS KGIDPPQVLSP 354

      Q:   672 DMVPPSERXXXXXXXXXXXXXXXXX EFTGVKELDDISQEIAQLQREKYSLEQDIREKEEAI RQ 851
      DMVPPSER                      EFTGVKELDDISQEIAQLQREKYSLEQDIREKEEAI+Q
25     S:   355 DMVPPSERGTPIPDSSSTLASGEFTGVKELDDISQEIAQLQREKYSLEQDIREKEEAIKQ 414

      Q:   852 KTSEVQELQNDLDR ETSSLQELEAQKQDAQDRLDEMDQQKAKLRDMLSDVRQKCQDETQM 1031
      KTSEVQELQNDLDR ETSSLQELEAQKQDAQDRLDEMDQQKAKLRDMLSDVRQKCQDETQ
30     S:   415 KTSEVQELQNDLDR ETSSLQELEAQKQDAQDRLDEMDQQKAKLRDMLSDVRQKCQDETQT 474

      Q:   1032 ISSLKTQIQSQESDLKSQEDDLNRAKSELNRLQQEETQLEQSIQAGRVQLETI I KSLKST 1211
      ISSLKTQIQSQESDLKSQEDDLNRAKSELNRLQQEETQLEQSIQAGR QLETI++SLK T
35     S:   475 ISSLKTQIQSQESDLKSQEDDLNRAKSELNRLQQEETQLEQSIQAGRAQLETIL RSLKCT 534

      Q:   1212 QDEINQARSKLSQLHESRQEAHRSLEQYDQVLDGAHGASLTDLANLSEGVSLAERGSFGA 1391
      QD+INQARSKLSQL ES  EAHRSLEQYDQV DG  G SL DLA L+EG+ LAERG FGA
40     S:   535 QDDINQARSKLSQLQESHLEAHRSLEQYDQVPDGVSGTSLPDLATLNEGILLAERGGFGA 594

      Q:   1392 MDDPFKNKALLFSNNTQELHPDPFQTEDPFKSDPFKGADPFKGDPFQNDPFAEQRQLNRS 1571
      MDDPFKNKALLFSNN+QELHPDPFQ EDPFKSDPFKGADPFKGDPFQ+DPF+EQ+
45     S:   595 MDDPFKNKALLFSNNSQELHPDPFQAE DPFKSDPFKGADPFKGDPFQSDPFSEQQTAATD 654

      Q:   1572 IWRG-PFQRK*P 1604
      + G PF+ P
      S:   655 PFGGDPFKESDP 666

```

The segment of gil968973 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 141. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities

are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
5 amino acid sequence set out in the sequence listing as SEQ ID NO. 142 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: NCI\_CGAP\_GCB1; Primary Dendritic Cells, lib 1 and to a  
10 lesser extent in Fetal Heart; neutrophils control; Jurkat Cells; Saos2, Dexamethosome Treated; NTERA2 teratocarcinoma cell line+retinoic acid (14 days); HL-60, PMA 4H, re-excision; Human Umbilical Vein, Reexcision; Stratagene neuroepithelium (#937231); Human Pancreas Tumor; Human Adipose; Bone Marrow Stromal Cell, untreated; Pancreas Islet Cell Tumor; Human Placenta; Early Stage Human Brain;  
15 Human Adult Pulmonary, re-excision; Endothelial cells-control; Activated T-cell(12h)/Thiouridine-re-excision; Hodgkin's Lymphoma II and Osteoblasts.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 102 as residues: Thr-17 to Cys-25, Pro-27 to Arg-41, Pro-55 to Thr-67, Ser-82 to Gln-91.

20 The tissue distribution in dendritic cells suggests the protein product of this clone is useful for the diagnosis and treatment of a variety of immune system disorders. Moreover, the expression of this gene product suggests a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells. This gene product may be involved in the  
25 regulation of cytokine production, antigen presentation, or other processes suggesting a usefulness in the treatment of cancer (e.g. by boosting immune responses).



Since the gene is expressed in cells of lymphoid origin, the natural gene product may be involved in immune functions. Therefore it may be also used as an agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosus, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, scleroderma and tissues.

Moreover, the protein may represent a secreted factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. In addition, this gene product may have commercial utility in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Moreover, the expression within cellular sources marked by proliferating cells, combined with the homology to the epidermal growth factor receptor, suggests this protein may play a role in the regulation of cellular division, and may show utility in the diagnosis, treatment, and/or prevention of developmental diseases and disorders, cancer, and other proliferative conditions. The protein is useful in modulating signal transduction pathways which may have utility in controlling cell growth, ameliorating cancer, etc.

Similarly, developmental tissues rely on decisions involving cell differentiation and/or apoptosis in pattern formation. Dysregulation of apoptosis can result in inappropriate suppression of cell death, as occurs in the development of some cancers, or in failure to control the extent of cell death, as is believed to occur in acquired immunodeficiency and certain neurodegenerative disorders, such as spinal

muscular atrophy (SMA). Therefore, the polynucleotides and polypeptides of the present invention are useful in treating, detecting, and/or preventing said disorders and conditions, in addition to other types of degenerative conditions. Thus this protein may modulate apoptosis or tissue differentiation and would be useful in the detection, treatment, and/or prevention of degenerative or proliferative conditions and diseases. The protein is useful in modulating the immune response to aberrant polypeptides, as may exist in proliferating and cancerous cells and tissues. The protein can also be used to gain new insight into the regulation of cellular growth and proliferation. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:52 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 4010 of SEQ ID NO:52, b is an integer of 15 to 4024, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:52, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 43**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: PERM TF274.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:53 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 549 of SEQ ID NO:53, b is an integer of 15 to 563, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:53, and where b is greater than or equal to a + 14.

#### FEATURES OF PROTEIN ENCODED BY GENE NO: 44

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gil289628 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "Caenorhabditis elegans protein". A partial alignment demonstrating the observed homology is shown immediately below.

```
>gi|289628 putative [Caenorhabditis elegans]
      Length = 383
```

Plus Strand HSPs:

```
Score = 184 (64.8 bits), Expect = 7.1e-11, P = 7.1e-11
Identities = 67/253 (26%), Positives = 103/253 (40%), Frame = +1
```

```
30  Q:  409 GIRGHSRWFVLVRVLLSLFIGAEIVAVHFSAEWFVGTVNTNTSYKAFSAARVTARVGLLV 588
      G+R  R    V  +L L +G  ++A      W  G+    T  ++  S  R+  A++G+  +
      S:  31 GVR-RKRVVTTVTYVLMMLAVGGALIASLIYPCWASGSQMIYTQFRGHSNERILAKIGVEI 89
```

5  
 Q: 589 GLEGINITLTGTPVHQLNETID-----YNEQFTWRXXXXXXXXXXXXXXXXXGLPDPVL 744  
 GL+ +N+TL + N+ + YNE F GLP P+L  
 S: 90 GLQKVNVTLKFERLLSSNDVLPGSDMTELYYNEGFDISGISSMAEALHHGLENGLPYPML 149

10  
 Q: 745 YLAEKFTPSSPCGLY-HQYHLAGHYASATLWVAFCFWLLSNVL-LSTPAPLYXXXXXXXXX 918  
 + E F+ + + Y +AGHY A +W AF W LS VL L P Y  
 S: 150 SVLEYFSLNQDSFDWGRHYRVAGHYTHAAIWFAFACWCLSVVLMFLPHNAYKSILATGI 209

15  
 Q: 919 XXXXXXXXXXXXXSISVPLCPLRLGSSA-----LTTQYGAAFWVTLATGVLCLFLPGA 1077  
 S P C LR+ + LT + F++ A G+LC+  
 S: 210 SCLIACLVYLLLS----P-CELRIAFTGENFERVDLTATFSFCFYLIFAIGILCVLCGLG 264

Q: 1078 VVSLQYVRPSALRTLDDQS 1134  
 + ++ R L T LD S  
 S: 265 LGICEHWRIYTLSTFLDAS 283

The segment of gil289628 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 143. Based on the structural similarity these homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 144 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Ulcerative Colitis; human tonsils.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:54 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention

are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1446 of SEQ ID NO:54, b is an integer of 15 to 1460, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:54, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 45**

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares fetal liver spleen 1NFLS; Ulcerative Colitis.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:55 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 996 of SEQ ID NO:55, b is an integer of 15 to 1010, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:55, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 46**

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gnllPIDld1026466 (all information available through the recited accession number is

incorporated herein by reference) which is described therein as "KIAA0610 protein [Homo sapiens]". A partial alignment demonstrating the observed homology is shown immediately below.

```
5      >gnl|PID|d1026466 (AB011182) KIAA0610 protein [Homo sapiens]
      Length = 686

      Plus Strand HSPs:

10     Score = 3153 (1109.9 bits), Expect = 0.0, Sum P(2) = 0.0
      Identities = 614/656 (93%), Positives = 614/656 (93%), Frame = +3

      Q:   93 EMEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGIGHLLRGISI 272
      EMEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGIGHLLRGISI
15     S:   20 EMEQEPQNGEPAEIKIIREAYKKAFLFVNKGLNTDELGQKEEAKNYYKQGIGHLLRGISI 79

      Q:   273 SSKSEHTGPGWESARQMQQMKKETLQNVTRTRLEILEKGLATSLQNDLQEVPKLYPEFPP 452
      SSKSEHTGPGWESARQMQQMKKETLQNVTRTRLEILEKGLATSLQNDLQEVPKLYPEFPP
20     S:   80 SSKSEHTGPGWESARQMQQMKKETLQNVTRTRLEILEKGLATSLQNDLQEVPKLYPEFPP 139

      Q:   453 KDMCEKLPEPQSFSSAPQHAEVNGNTXXXXXXXXXXXXXXXXXXXXEAPPAYTPQAA 632
      KDMCEKLPEPQSFSSAPQHAEVNGNT                                EAPPAYTPQAA
      S:   140 KDMCEKLPEPQSFSSAPQHAEVNGNTSTPSAGAVAAPASLSLPSQSCPAEAPPAYTPQAA 199

25     Q:   633 EGHYTVSYGTDSGEFSSVGEEFYRNHSQPPPLETLGLDADELILIPNGVQIFFVNPAGEV 812
      EGHYTVSYGTDSGEFSSVGEEFYRNHSQPPPLETLGLDADELILIPNGVQIFFVNPAGEV
      S:   200 EGHYTVSYGTDSGEFSSVGEEFYRNHSQPPPLETLGLDADELILIPNGVQIFFVNPAGEV 259

30     Q:   813 SAPSYPGYLRIVRFLDNSLDTVLNRPPGFLQVCDWLYPLVPDRSPVLKCTAGAYMFPDTM 992
      SAPSYPGYLRIVRFLDNSLDTVLNRPPGFLQVCDWLYPLVPDRSPVLKCTAGAYMFPDTM
      S:   260 SAPSYPGYLRIVRFLDNSLDTVLNRPPGFLQVCDWLYPLVPDRSPVLKCTAGAYMFPDTM 319

      Q:   993 LQAAGCFVGVVLSSSELPEDDRELFE DLLRQMSDLRLQANWNRAEEENE FQIPGRTRPSSD 1172
      LQAAGCFVGVVLSSSELPEDDRELFE DLLRQMSDLRLQANWNRAEEENE FQIPGRTRPSSD
35     S:   320 LQAAGCFVGVVLSSSELPEDDRELFE DLLRQMSDLRLQANWNRAEEENE FQIPGRTRPSSD 379

      Q:   1173 QLKEASGTDVKQLDQGNKDVRHKGRGKRAKDTSSSEVNLSHIVXXXXXXXXXXXXXXXXX 1352
      QLKEASGTDVKQLDQGNKDVRHKGRGKRAKDTSSSEVNLSHIV
      S:   380 QLKEASGTDVKQLDQGNKDVRHKGRGKRAKDTSSSEVNLSHIVPCEPVPEEKPKELPEW 439

40     Q:   1353 XXXVAHNILSGASWVSWGLVKGAETGKAIQKGASKLRERIQPEEKPEVSPAVTKGLYI 1532
      VAHNILSGASWVSWGLVKGAETGKAIQKGASKLRERIQPEEKPEVSPAVTKGLYI
      S:   440 SEKVAHNILSGASWVSWGLVKGAETGKAIQKGASKLRERIQPEEKPEVSPAVTKGLYI 499

45     Q:   1533 AKQATGGAAKVSQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAM 1712
      AKQATGGAAKVSQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAM
      S:   500 AKQATGGAAKVSQFLVDGVCTVANCVGKELAPHVKKHGSKLVPESLKKDKDGKSPLDGAM 559

50     Q:   1713 VVAASSVQGFSTVWQGLECAAKCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGV 1892
      VVAASSVQGFSTVWQGLECAAKCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGV
      S:   560 VVAASSVQGFSTVWQGLECAAKCIVNNVSAETVQTVRYKYGYNAGEATHHAVDSAVNVGV 619

      Q:   1893 TAYNINNIGIKAMVKKTTATQTGHTLLEDYQIVDNSQRENQEGAANVNVVRGEKDEQT 2060
```

TAYNINNIGIKAMVKKATATGTGHTLLEDYQIVDNSQRENQEGAANVNVVRGEKDEQT  
 S: 620 TAYNINNIGIKAMVKKATATGTGHTLLEDYQIVDNSQRENQEGAANVNVVRGEKDEQT 675

The segment of gnl|PID|d1026466 that is shown as "S" above is set out in the  
 5 sequence listing as SEQ ID NO. 145. Based on the structural similarity these  
 homologous polypeptides are expected to share at least some biological activities.  
 Such activities are known in the art, some of which are described elsewhere herein.  
 Assays for determining such activities are also known in the art, some of which have  
 been described elsewhere herein.

10 Preferred polypeptides of the invention comprise a polypeptide having the  
 amino acid sequence set out in the sequence listing as SEQ ID NO. 146 which  
 corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a  
 sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following  
 15 tissues/cDNA libraries: Soares retina N2b4HR and to a lesser extent in Glioblastoma;  
 Human 8 Week Whole Embryo; Nine Week Old Early Stage Human; Synovial  
 Fibroblasts (IL1/TNF), sub; Soares placenta Nb2HP; Morton Fetal Cochlea; Human  
 Lung; Human Bone Marrow, re-excision; Stratagene fetal spleen (#937205); Soares  
 breast 2NbHBst; NTERA2, control; Soares breast 3NbHBst; Dendritic cells, pooled;  
 20 Endothelial cells-control; Primary Dendritic Cells, lib 1; Smooth muscle-edited A;  
 Soares ovary tumor NbHOT; NTERA2 teratocarcinoma cell line+retinoic acid (14  
 days); Human Lung Cancer, re-excision; Pancreas Tumor PCA4 Tu; Synovial IL-  
 1/TNF stimulated; pBMC stimulated w/ poly I/C; Smooth muscle, IL1b induced; H  
 Female Bladder, Adult; NTERA2 + retinoic acid, 14 days; Human Adipose Tissue,  
 25 re-excision; Human Ovary; Human Manic Depression Tissue; Hippocampus,  
 Alzheimer Subtracted; Human Osteoblasts II; 12 Week Old Early Stage Human, II;  
 Human Umbilical Vein Endothelial Cells, uninduced; T-Cell PHA 24 hrs; Human

Rhabdomyosarcoma; Epithelial-TNF $\alpha$  and INF induced; Bone Marrow Stromal Cell, untreated; Resting T-Cell Library,II; Human Placenta; Adipocytes; Soares\_multiple\_sclerosis\_2NbHMSP; Human Placenta; Human Adult Pulmonary, re-excision; Anergic T-cell; Smooth muscle,control; Human Testes; Hodgkin's

5 Lymphoma II and Human Cerebellum.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:56 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically

10 excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2913 of SEQ ID NO:56, b is an integer of 15 to 2927, where both a and b correspond to the positions of

15 nucleotide residues shown in SEQ ID NO:56, and where b is greater than or equal to a + 14.

#### **FEATURES OF PROTEIN ENCODED BY GENE NO: 47**

The computer algorithm BLASTX has been used to determine that the

20 translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. gnllPIDle218082 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "head-elevated expression in 0.9 kb [Drosophila melanogaster]". A partial alignment demonstrating

25 the observed homology is shown immediately below.

```
>gnl|PID|e218082 head-elevated expression in 0.9 kb [Drosophila  
melanogaster]
```



>sp|Q24269|Q24269 CODING REGION CLONED BY GAL4 ENHANCER TRAP  
LINE.

Length = 223

5 Plus Strand HSPs:

Score = 155 (54.6 bits), Expect = 1.2e-08, P = 1.2e-08

Identities = 35/104 (33%), Positives = 58/104 (55%), Frame = +2

10 Q: 521 AKHLARQLEEKDRVLKKQDAFYKEQLARLEERSSEFYRVTTTEQYQKAAEEVEAKF-KRYE 697  
A + RQ+ ++ L K D +KE++A+LEE + + ++Y A E V +F  
S: 100 AADVQRQMNQE---LIKNDLWKERMAKLEENLKKTNLTILEKEYANAVENVHKRFVSTAS 156

15 Q: 698 SH--PVCADLQAKILQCYRENTHTLTKCSALATQYMHCVN-HAKQSM 829  
SH P C DL++++L CYR + +TLKC Q+ C++ H Q +  
S: 157 SHKVPPCQDLKSQLLACYRAHPGETLKCMEEVAQFRQCIDLHRVQKL 203

The segment of gnlPIDle218082 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 147. Based on the structural similarity these  
20 homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
25 amino acid sequence set out in the sequence listing as SEQ ID NO. 148 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares melanocyte 2NbHM and to a lesser extent in Soares  
30 fetal liver spleen 1NFLS; Soares breast 3NbHBst; Keratinocyte; Human Microvascular Endothelial Cells, fract. A; Bone Marrow Cell Line (RS4,11); Supt Cells, cyclohexamide treated; Stratagene colon (#937204); Ovarian Tumor 10-3-95; Jurkat T-Cell, S phase; Breast Cancer Cell line, angiogenic; Human umbilical vein endothelial cells, IL-4 induced; Human Gall Bladder; Human Adult Pulmonary, re-  
35 excision; Osteoblasts; Soares\_pregnant\_uterus\_NbHPU; Stratagene NT2 neuronal

- precursor 937230; Soares\_fetal\_heart\_NbHH19W; Human Adult Spleen; Hodgkin's Lymphoma I; Human Fetal Bone; human colon cancer; Smooth muscle-ILb induced; Human Lung; Human Adult Heart, re-excision; Cem cells cyclohexamide treated; Healing groin wound - zero hr post-incision (control); STROMAL -
- 5 OSTEOCLASTOMA; Soares\_pregnant\_uterus\_NbHPU; Human endometrial stromal cells-treated with estradiol; Human Osteoclastoma, re-excision; Myeloid Progenitor Cell Line; Human Adult Small Intestine; KMH2; Human Prostate; H. Kidney Medulla, re-excision; Human Pancreas Tumor; Spinal cord; Human Thymus Stromal Cells; Human Adrenal Gland Tumor; NTERA2, control; Pancreas Islet Cell Tumor;
- 10 12 Week Old Early Stage Human; Human T-Cell Lymphoma; Neutrophils IL-1 and LPS induced; Smooth muscle, serum treated; Adipocytes; Human Testes Tumor; Soares\_fetal\_lung\_NbHL19W; Bone marrow; Human Fetal Heart; Activated T-Cell (12hs)/Thiouridine labelledEco; T Cell helper I; T cell helper II; Human 8 Week Whole Embryo; Nine Week Old Early Stage Human; Human Cerebellum;
- 15 Soares\_fetal\_lung\_NbHL19W; Soares\_senescent\_fibroblasts\_NbHSF; Primary Dendritic Cells, lib 1 and Soares infant brain 1NIB.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:57 and may have been publicly available prior to conception of

20 the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1634 of SEQ ID NO:57, b

25 is an integer of 15 to 1648, where both a and b correspond to the positions of

nucleotide residues shown in SEQ ID NO:57, and where b is greater than or equal to a + 14.

## FEATURES OF PROTEIN ENCODED BY GENE NO: 48

5           The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. dbj|AB000713\_1 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "CPE-receptor [Mus musculus]". A partial alignment demonstrating the observed homology is shown immediately below.

```

15      >dbj|AB000713_1 (AB000713) CPE-receptor [Mus musculus]
          Length = 210

      Plus Strand HSPs:

      Score = 573 (201.7 bits), Expect = 5.7e-55, P = 5.7e-55
      Identities = 104/208 (50%), Positives = 147/208 (70%), Frame = +3
20
      Q:   87 MANSGLQLLGFSMALLGWVGLVACTAIPQWQMSSYAGDNIITAQAMYKGLWMDCVTQSTG 266
          MA+ GLQ+LG S+A+LGW+G++ A+P W+++++ G NI+TAQ ++GLWM+CV QSTG
      S:   1 MASMGLQVLGISLAVLGWLGIILSCALPMWRVTAFIGSNIVTAQTSWEGLWMNCVVQSTG 60

25      Q:  267 MMSCKMYDSVLALSAALQATRALMVVSLVLGFLAMFVATMGMKCTRCGGDDKVKKARIAM 446
          M CKMYDS+LAL LQA RALMV+S+++G L M ++ +G KCT C D+ VK A+I +
      S:   61 QMQCKMYDSMLALPQDLQAARALMVISIIVGALGMLLSVVGKCTNCMEDETVK-AKIMI 119

30      Q:  447 GGGIIFIVAGLAALVACSWYGHQIVTDFYNPLIPTNIKYEFGPAIFIGWAGSALVILGGA 626
          G +FIVA + +V SW H ++ DFYNP++ + K E G ++++GWA S L++LGG
      S:  120 TAGAVFIVASMLIMVPVSWTAHNVIRDFYNPMVASGQKREMGASLYVGWAASGLLLLGGG 179

      Q:   627 LLSCSCP--GNES--KAGYRAPRSYPKSN 701
          LL CSCP N+ A Y A RS P SN
35      S:  180 LLCCSCPPRSNDKPYSAKYSAARSVPASN 208
  
```

The segment of dbj|AB000713\_1 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 149. Based on the structural similarity these

homologous polypeptides are expected to share at least some biological activities. Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

5 Preferred polypeptides of the invention comprise a polypeptide having the amino acid sequence set out in the sequence listing as SEQ ID NO. 150 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following  
10 tissues/cDNA libraries: Stratagene colon (#937204) and to a lesser extent in Soares ovary tumor NbHOT; Human Epididymus; Human Prostate; Ulcerative Colitis; Epithelial-TNF $\alpha$  and INF induced; Colon Carcinoma; Colon Normal III; Keratinocyte; Soares fetal liver spleen 1NFLS; Salivary Gland; Human Fetal Kidney; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Human Uterine Cancer; Soares breast  
15 2NbHBst; Human Endometrial Tumor; Soares\_fetal\_lung\_NbHL19W; Human Cornea, subtracted; Human Umbilical Vein Endothelial Cells, fract. B; Soares\_total\_fetus\_Nb2HF8\_9w; Colon, normal; Human Colon Cancer, subtracted; LNCAP untreated; Human Colon; H. Epididymus, caput & corpus; Human adult small intestine, re-excision; Human Colon Cancer, re-excision; Stratagene placenta  
20 (#937225); Amniotic Cells - Primary Culture; Human Stomach, re-excision; Soares\_placenta\_8to9weeks\_2NbHP8to9W; Ovarian Tumor 10-3-95; H. Kidney Medulla, re-excision; human ovarian cancer; Human Fetal Dura Mater; Human Primary Breast Cancer Reexcision; Human Pancreas Tumor; Merkel Cells; Soares\_NhHMPu\_S1; Soares\_fetal\_heart\_NbHH19W; Hepatocellular Tumor, re-  
25 excision; Human Gall Bladder; Colon Tumor; Colon Normal II and Hodgkin's Lymphoma II.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 108 as residues: Arg-106 to Lys-114, Pro-187 to Gly-194, Ala-197 to Val-211.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:58 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 1000 of SEQ ID NO:58, b is an integer of 15 to 1014, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:58, and where b is greater than or equal to a + 14.

## 15 **FEATURES OF PROTEIN ENCODED BY GENE NO: 49**

It has been discovered that this gene is expressed primarily in H. Ovarian Tumor, II, OV5232.

Preferred epitopes include those comprising a sequence shown in SEQ ID NO. 109 as residues: Ser-30 to Ser-36.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:59 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the

general formula of a-b, where a is any integer between 1 to 1022 of SEQ ID NO:59, b is an integer of 15 to 1036, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:59, and where b is greater than or equal to a + 14.

5

## FEATURES OF PROTEIN ENCODED BY GENE NO: 50

The computer algorithm BLASTX has been used to determine that the translation product of this gene shares sequence homology with, as a non-limiting example, the sequence accessible through the following database accession no. 10 gil2062700 (all information available through the recited accession number is incorporated herein by reference) which is described therein as "non-histone chromosomal protein [Homo sapiens]". A partial alignment demonstrating the observed homology is shown immediately below.

```

15      >gi|2062700 non-histone chromosomal protein [Homo sapiens]
      >sp|O00479|O00479
          NON-HISTONE CHROMOSOMAL PROTEIN.
          Length = 90

20      Plus Strand HSPs:

      Score = 166 (58.4 bits), Expect = 8.6e-11, P = 8.6e-11
      Identities = 39/71 (54%), Positives = 39/71 (54%), Frame = +3

25      Q:   273 EPQRRSARLSXXXXXXXXXXXXXXXXXXXXGKLPXXXXXXXXXXXXXXXXNNPAKNRDASTLQ 452
          EPQRRSARLS                                GKLP                                NNPAKNRDASTLQ
      S:   20  EPQRRSARLSAKPAPPKPEPRPKKASAKKGKLPKGRKGKADAGKDGNNPAKNRDASTLQ 79

      Q:   453 SQKAEGTGDAK 485
          SQKAEGTGDAK
30      S:   80 SQKAEGTGDAK 90

```

The segment of gil2062700 that is shown as "S" above is set out in the sequence listing as SEQ ID NO. 151. Based on the structural similarity these  
35 homologous polypeptides are expected to share at least some biological activities.

Such activities are known in the art, some of which are described elsewhere herein. Assays for determining such activities are also known in the art, some of which have been described elsewhere herein.

Preferred polypeptides of the invention comprise a polypeptide having the  
5 amino acid sequence set out in the sequence listing as SEQ ID NO. 152 which corresponds to the "Q" sequence in the alignment shown above (gaps introduced in a sequence by the computer are, of course, removed).

It has been discovered that this gene is expressed primarily in the following tissues/cDNA libraries: Soares melanocyte 2NbHM and to a lesser extent in  
10 Soares\_multiple\_sclerosis\_2NbHMSP; Soares infant brain 1NIB;  
Soares\_parathyroid\_tumor\_NbHPA; Human Adult Pulmonary, re-excision; Spleen, Chronic lymphocytic leukemia; B Cell lymphoma; Stratagene colon (#937204); Stratagene hNT neuron (#937233); 12 Week Old Early Stage Human;  
Soares\_senescent\_fibroblasts\_NbHSF; human tonsils;  
15 Soares\_fetal\_heart\_NbHH19W; Human Microvascular Endothelial Cells, fract. A; Human 8 Week Whole Embryo; Nine Week Old Early Stage Human; Primary Dendritic Cells, lib 1; Normal lung; Supt cells, cyclohexamide treated, differentially expressed; Kidney Pyramids; Thymus; Human Kidney; HUMAN TONSILS, FRACTION 2; Human Thyroid; Human T-cell lymphoma, re-excision; HSA 172  
20 Cells; Human adult small intestine, re-excision; Cem cells cyclohexamide treated; Soares\_multiple\_sclerosis\_2NbHMSP; Human Tonsils, Lib 2; Apoptotic T-cell, re-excision; Stratagene fibroblast (#937212); HEL cell line; Human Colon Cancer, re-excision; Synovial IL-1/TNF stimulated; Ovarian Tumor 10-3-95; Human Colon, re-excision; TF-1 Cell Line GM-CSF Treated; L428; 12 Week Old Early Stage Human,  
25 II; Human Activated T-Cells; T-Cell PHA 24 hrs; Stromal cell TF274; Human Ovarian Cancer Reexcision; Human Thymus Stromal Cells; Rejected Kidney, lib 4;

Pancreas Islet Cell Tumor; Colon Tumor; Human T-Cell Lymphoma; Colon Carcinoma; Adipocytes; H Macrophage (GM-CSF treated), re-excision; Soares\_fetal\_lung\_NbHL19W; Human Testes, Reexcision; Human Fetal Heart; Human Neutrophil, Activated; Human Osteoclastoma; Monocyte activated; HUMAN  
5 B CELL LYMPHOMA; Human Bone Marrow, treated; Activated T-cell(12h)/Thiouridine-re-excision; Human Testes; Hodgkin's Lymphoma II; Osteoblasts; Keratinocyte; Human Cerebellum and Soares placenta Nb2HP.

Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are  
10 related to SEQ ID NO:60 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides comprising a nucleotide sequence described by the  
15 general formula of a-b, where a is any integer between 1 to 1952 of SEQ ID NO:60, b is an integer of 15 to 1966, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:60, and where b is greater than or equal to a + 14.



Table 1

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of 5' NT of Clone Seq. Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
1	HFXDR47	203959 4/26/99	Lambda ZAP II	11	1808	1	1808	207	207	61	1	34	35	78
2	HT5GR59	203959 4/26/99	Uni-ZAP XR	12	1743	1	1743	135	135	62	1	23	24	31
3	HTHCA30	203959 4/26/99	Uni-ZAP XR	13	627	1	627	158	158	63	1			23

Gene No.	cDNA Clone ID	ATCC Deposit No.:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of AA of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Secreted Portion	Last AA of ORF
4	HTHDV50	203959 4/26/99	Uni-ZAP XR	14	535	1	391		119	64	1		23
5	HUSGB23	203959 4/26/99	pSport1	15	719	1	718	168	168	65	1	20	37
6	HNHHB10	203959 4/26/99	Uni-ZAP XR	16	901	1	901	215	215	66	1	30	59
7	HPIAA80	203959 4/26/99	Uni-ZAP XR	17	919	312	919		314	67	1	14	37

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of 5' NT of AA of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
8	HPTRI42	203959 4/26/99	pBluescript	18	1098	1	1098		266	68	1	19	20	71
9	HTWCE14	203959 4/26/99	pSport1	19	1054	828	1036	892	892	69	1	29	30	54
10	HTWBZ57	203959 4/26/99	pSport1	20	1667	777	1658	990	990	70	1	28	29	28
11	HCFBV39	203959 4/26/99	pSport1	21	1759	890	1759		419	71	1			3

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of AA of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
12	HPTRQ52	203959 4/26/99	pBluescript	22	658	1	658	224	224	72	1	15	16	33
13	HPTVH24	203959 4/26/99	pBluescript	23	1463	576	1453	766	766	73	1	19	20	35
14	HPTVH59	203959 4/26/99	pBluescript	24	475	1	475		144	74	1	20	21	44
15	HPTVI96	203959 4/26/99	pBluescript	25	520	1	520	271	271	75	1			6

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of AA of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
16	HUSGU40	203959 4/26/99	pSport1	26	1054	1	1054		500	76	1	22	23	46
17	HUSYG26	203959 4/26/99	pSport1	27	744	363	744	421	421	77	1	29	30	69
18	HOVCJ71	203959 4/26/99	pSport1	28	1463	1	1463		147	78	1	18	19	46
19	HSAYC41	203959 4/26/99	Uni-ZAP XR	29	214	1	214	106	106	79	1	16	17	36

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Clone Seq.	Start Codon	Signal Pep	AA of ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
20	HSKDS47	203959 4/26/99	Uni-ZAP XR	30	1159	474	1159	740			80	1	14	15	17
21	HSAYR62	203959 4/26/99	Uni-ZAP XR	31	3116	1559	2971	1900			81	1			23
22	HNTMH79	203959 4/26/99	pSport1	32	922	1	922	48	48		82	1	35	36	38
23	HSKYR49	203959 4/26/99	pBluescript	33	2631	1750	2631	2066	2066		83	1	20	21	81

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of Clone Seq.	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
24	HOEEZ62	203959 4/26/99	Uni-ZAP XR	34	385	1	385	258	258	84	1			17
25	HUFAU71	203959 4/26/99	pSport1	35	2146	887	2130	1041	1041	85	1			31
26	HUFBK95	203959 4/26/99	pSport1	36	1384	1	1384	174	174	86	1			16
27	HUFAO94	203959 4/26/99	pSport1	37	739	1	739	229	229	87	1			13

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
28	HUFBV62	203959 4/26/99	pSport1	38	743	1	743	453	453	88	1	23	24	31
29	HTWEG06	203959 4/26/99	pSport1	39	1066	1	1066		132	89	1	15	16	57
30	HSDJF04	203959 4/26/99	Uni-ZAP XR	40	1144	1	1144		74	90	1	29	30	116
31	HTXJY08	203959 4/26/99	Uni-ZAP XR	41	1187	12	1187	108	108	91	1			16



Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of 5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Secreted Portion	Last AA of ORF
32	HPQBT17	203959 4/26/99	Lambda ZAP II	42	1220	3	1208	236	236	92	1		12
33	HPQAN50	203959 4/26/99	Lambda ZAP II	43	1285	1	1228	164	164	93	1	22	44
34	HPQAO80	203959 4/26/99	Lambda ZAP II	44	985	10	983	284	284	94	1		2
35	H6EDM64	203959 4/26/99	Uni-ZAP XR	45	2610	1275	2377	1448	1448	95	1		6

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of 5' NT of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Secreted Portion	Last AA of ORF	
36	HPQCF94	203959 4/26/99	Lambda ZAP II	46	951	8	944	212	212	96	1	23	24	28
37	HSDZM54	203959 4/26/99	pBluescript	47	554	1	554	445	445	97	1	15	16	36
38	HMIAG72	203959 4/26/99	Uni-ZAP XR	48	3052	597	3052	740	740	98	1	22	23	38
39	HISAT67	203959 4/26/99	pSport1	49	2154	1061	2142	1239	1239	99	1	35	36	56

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
40	HT5FX76	203959 4/26/99	Uni-ZAP XR	50	643	1	643	98	98	100	1	34	35	52
41	HT5FX79	203959 4/26/99	Uni-ZAP XR	51	682	59	682		228	101	1	23	24	50
42	HNTRQ40	203959 4/26/99	pSport1	52	4024	1	4024	132	132	102	1	16	17	91
43	HPQRS74	203959 4/26/99	Lambda ZAP II	53	563	1	563	146	146	103	1	35	36	36

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	3' NT of Clone Seq.	5' NT of Start Codon	5' NT of AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
44	HUFCL31	203959 4/26/99	pSport1	54	1460	1	1460		287	104	1			26
45	HUFEF79	203959 4/26/99	pSport1	55	1010	1	1010	201	201	105	1			20
46	HOUFS04	203959 4/26/99	Uni-ZAP XR	56	2927	457	2882	520	520	106	1	42	43	72
47	HOGAH40	203959 4/26/99	pCMV Sport 2.0	57	1648	1058	1648	1339	1339	107	1	23	24	25

Gene No.	cDNA Clone ID	ATCC Deposit No:Z and Date	Vector	NT SEQ ID NO: X	Total NT Seq.	5' NT of Clone Seq.	5' NT 3' NT of Clone Seq.	5' NT of AA of Start Codon	5' NT of First AA of Signal Pep	AA SEQ ID NO: Y	First AA of Sig Pep	Last AA of Sig Pep	First AA of Secreted Portion	Last AA of ORF
48	HOGAR71	203959 4/26/99	pCMVSPORT 2.0	58	1014	1	1014	87	87	108	1	24	25	211
49	HOFNB74	203959 4/26/99	pCMVSPORT 2.0	59	1036	1	1036	138	138	109	1	34	35	39
50	HMEJK28	203959 4/26/99	Lambda ZAP II	60	1966	710	1933	943	943	110	1			26

Table 1 summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:X" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually three to five overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:X.

The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." Some of the deposits contain multiple different clones corresponding to the same gene. "Vector" refers to the type of vector contained in the cDNA Clone ID.

"Total NT Seq." refers to the total number of nucleotides in the contig identified by "Gene No." The deposited clone may contain all or most of these sequences, reflected by the nucleotide position indicated as "5' NT of Clone Seq." and the "3' NT of Clone Seq." of SEQ ID NO:X. The nucleotide position of SEQ ID NO:X of the putative start codon (methionine) is identified as "5' NT of Start Codon." Similarly, the nucleotide position of SEQ ID NO:X of the predicted signal sequence is identified as "5' NT of First AA of Signal Pep."

The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:Y," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

The first and last amino acid position of SEQ ID NO:Y of the predicted signal peptide is identified as "First AA of Sig Pep" and "Last AA of Sig Pep." The predicted first amino acid position of SEQ ID NO:Y of the secreted portion is

identified as "Predicted First AA of Secreted Portion." Finally, the amino acid position of SEQ ID NO:Y of the last amino acid in the open reading frame is identified as "Last AA of ORF."

SEQ ID NO:X (where X may be any of the polynucleotide sequences disclosed in the sequence listing) and the translated SEQ ID NO:Y (where Y may be any of the polypeptide sequences disclosed in the sequence listing) are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further below. For instance, SEQ ID NO:X is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:X or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:Y may be used, for example, to generate antibodies which bind specifically to proteins containing the polypeptides and the secreted proteins encoded by the cDNA clones identified in Table 1.

Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:X and the predicted

translated amino acid sequence identified as SEQ ID NO:Y, but also a sample of plasmid DNA containing a human cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods.

5 The predicted amino acid sequence can then be verified from such deposits.

Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited human cDNA, collecting the protein, and determining its sequence.

10 The present invention also relates to the genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate  
15 sources of genomic material.

Also provided in the present invention are allelic variants, orthologs, and/or species homologs. Procedures known in the art can be used to obtain full-length genes, allelic variants, splice variants, full-length coding portions, orthologs, and/or species homologs of genes corresponding to SEQ ID NO:X, SEQ ID NO:Y, or a  
20 deposited clone, using information from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

25 The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly



produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

The polypeptides may be in the form of the secreted protein, including the  
5 mature form, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification , such as multiple histidine residues, or an additional sequence for stability during recombinant production.

10 The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, including the secreted polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40  
15 (1988). Polypeptides of the invention also can be purified from natural, synthetic or recombinant sources using techniques described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the secreted protein.

The present invention provides a polynucleotide comprising, or alternatively consisting of, the nucleic acid sequence of SEQ ID NO:X, and/or a cDNA contained  
20 in ATCC deposit Z. The present invention also provides a polypeptide comprising, or alternatively, consisting of, the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide encoded by the cDNA contained in ATCC deposit Z. Polynucleotides encoding a polypeptide comprising, or alternatively consisting of the polypeptide sequence of SEQ ID NO:Y and/or a polypeptide sequence encoded by the cDNA  
25 contained in ATCC deposit Z are also encompassed by the invention.

#### **Signal Sequences**

The present invention also encompasses mature forms of the polypeptide having the polypeptide sequence of SEQ ID NO:Y and/or the polypeptide sequence encoded by the cDNA in a deposited clone. Polynucleotides encoding the mature forms (such as, for example, the polynucleotide sequence in SEQ ID NO:X and/or the polynucleotide sequence contained in the cDNA of a deposited clone) are also encompassed by the invention. According to the signal hypothesis, proteins secreted by mammalian cells have a signal or secretory leader sequence which is cleaved from the mature protein once export of the growing protein chain across the rough endoplasmic reticulum has been initiated. Most mammalian cells and even insect cells cleave secreted proteins with the same specificity. However, in some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide.

Methods for predicting whether a protein has a signal sequence, as well as the cleavage point for that sequence, are available. For instance, the method of McGeoch, *Virus Res.* 3:271-286 (1985), uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje, *Nucleic Acids Res.* 14:4683-4690 (1986) uses the information from the residues surrounding the cleavage site, typically residues -13 to +2, where +1 indicates the amino terminus of the secreted protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80%. (von Heinje, *supra.*) However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

In the present case, the deduced amino acid sequence of the secreted polypeptide was analyzed by a computer program called SignalP (Henrik Nielsen et al., Protein Engineering 10:1-6 (1997)), which predicts the cellular location of a protein based on the amino acid sequence. As part of this computational prediction of localization, the methods of McGeoch and von Heinje are incorporated. The analysis of the amino acid sequences of the secreted proteins described herein by this program provided the results shown in Table 1.

As one of ordinary skill would appreciate, however, cleavage sites sometimes vary from organism to organism and cannot be predicted with absolute certainty. Accordingly, the present invention provides secreted polypeptides having a sequence shown in SEQ ID NO:Y which have an N-terminus beginning within 5 residues (i.e., + or - 5 residues) of the predicted cleavage point. Similarly, it is also recognized that in some cases, cleavage of the signal sequence from a secreted protein is not entirely uniform, resulting in more than one secreted species. These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

Moreover, the signal sequence identified by the above analysis may not necessarily predict the naturally occurring signal sequence. For example, the naturally occurring signal sequence may be further upstream from the predicted signal sequence. However, it is likely that the predicted signal sequence will be capable of directing the secreted protein to the ER. Nonetheless, the present invention provides the mature protein produced by expression of the polynucleotide sequence of SEQ ID NO:X and/or the polynucleotide sequence contained in the cDNA of a deposited clone, in a mammalian cell (e.g., COS cells, as described below). These polypeptides, and the polynucleotides encoding such polypeptides, are contemplated by the present invention.

**Polynucleotide and Polypeptide Variants**

The present invention is directed to variants of the polynucleotide sequence disclosed in SEQ ID NO:X, the complementary strand thereto, and/or the cDNA  
5 sequence contained in a deposited clone.

The present invention also encompasses variants of the polypeptide sequence disclosed in SEQ ID NO:Y and/or encoded by a deposited clone.

"Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential  
10 properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

The present invention is also directed to nucleic acid molecules which comprise, or alternatively consist of, a nucleotide sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for example, the nucleotide  
15 coding sequence in SEQ ID NO:X or the complementary strand thereto, the nucleotide coding sequence contained in a deposited cDNA clone or the complementary strand thereto, a nucleotide sequence encoding the polypeptide of SEQ ID NO:Y, a nucleotide sequence encoding the polypeptide encoded by the cDNA contained in a deposited clone, and/or polynucleotide fragments of any of  
20 these nucleic acid molecules (e.g., those fragments described herein).

Polynucleotides which hybridize to these nucleic acid molecules under stringent hybridization conditions or lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

The present invention is also directed to polypeptides which comprise, or  
25 alternatively consist of, an amino acid sequence which is at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% identical to, for example, the polypeptide sequence

shown in SEQ ID NO:Y, the polypeptide sequence encoded by the cDNA contained in a deposited clone, and/or polypeptide fragments of any of these polypeptides (e.g., those fragments described herein).

By a nucleic acid having a nucleotide sequence at least, for example, 95%  
5 "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the nucleic acid is identical to the reference sequence except that the nucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a nucleic acid having a nucleotide sequence at least 95%  
10 identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified  
15 as described herein.

As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best  
20 overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245(1990)). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by  
25 converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to

calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

5           If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent  
10 identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using  
15 the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

20           For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number  
25 of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly

matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to made for the purposes of the present invention.

By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular polypeptide is at least 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, an amino acid sequences shown in Table 1 (SEQ ID NO:Y) or to the amino acid sequence encoded by cDNA contained in a deposited clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject

sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. 6:237-245(1990)). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said

5 global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

10 If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the

15 query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from

20 the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity

25 score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.



For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10  
5 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100  
10 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not  
15 matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter  
20 the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host  
25 (change codons in the human mRNA to those preferred by a bacterial host such as *E. coli*).

Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the secreted protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein. (Dobeli et al., J. Biotechnology 7:199-216 (1988).)

Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." (See, Abstract.) In fact, only 23 unique amino acid sequences, out of more than 3,500

nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the secreted form will likely be retained when less than the majority of the residues of the secreted form are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

Thus, the invention further includes polypeptide variants which show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used.

5 (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the  
10 protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved. Moreover, tolerated conservative amino acid substitutions involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic  
15 residues Asp and Glu; replacement of the amide residues Asn and Gln, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

Besides conservative amino acid substitution, variants of the present invention include (i) substitutions with one or more of the non-conserved amino acid residues,  
20 where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino  
25 acids, such as, for example, an IgG Fc fusion region peptide, or leader or secretory

sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

10 A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and  
15 still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or  
20 fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5, 5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

25 **Polynucleotide and Polypeptide Fragments**

The present invention is also directed to polynucleotide fragments of the polynucleotides of the invention.

In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence which: is a portion of that contained in  
5 a deposited clone, or encoding the polypeptide encoded by the cDNA in a deposited clone; is a portion of that shown in SEQ ID NO:X or the complementary strand thereto, or is a portion of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:Y. The nucleotide fragments of the invention are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt,  
10 and even more preferably, at least about 40 nt, at least about 50 nt, at least about 75 nt, or at least about 150 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in a deposited clone or the nucleotide sequence shown in SEQ ID NO:X. In this context "about" includes the particularly recited value, a value larger  
15 or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. These nucleotide fragments have uses that include, but are not limited to, as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

Moreover, representative examples of polynucleotide fragments of the  
20 invention, include, for example, fragments comprising, or alternatively consisting of, a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500,  
25 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:X, or the

complementary strand thereto, or the cDNA contained in a deposited clone. In this context "about" includes the particularly recited ranges, and ranges larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini.

Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein. Polynucleotides which hybridize to these nucleic acid molecules under stringent hybridization conditions or lower stringency conditions are also encompassed by the invention, as are polypeptides encoded by these polynucleotides.

In the present invention, a "polypeptide fragment" refers to an amino acid sequence which is a portion of that contained in SEQ ID NO:Y or encoded by the cDNA contained in a deposited clone. Protein (polypeptide) fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges or values, and ranges or values larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes. Polynucleotides encoding these polypeptides are also encompassed by the invention.

Preferred polypeptide fragments include the secreted protein as well as the mature form. Further preferred polypeptide fragments include the secreted protein or the mature form having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of either the secreted polypeptide or the

mature form. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the secreted protein or mature form.

Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotides encoding these polypeptide fragments are also preferred.

Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:Y falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotides encoding these domains are also contemplated.

Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity. Polynucleotides encoding these polypeptide fragments are also encompassed by the invention.

Preferably, the polynucleotide fragments of the invention encode a polypeptide which demonstrates a functional activity. By a polypeptide demonstrating a "functional activity" is meant, a polypeptide capable of displaying one or more known functional activities associated with a full-length (complete) polypeptide of invention protein. Such functional activities include, but are not limited to, biological activity, antigenicity [ability to bind (or compete with a



polypeptide of the invention for binding) to an antibody to the polypeptide of the invention], immunogenicity (ability to generate antibody which binds to a polypeptide of the invention), ability to form multimers with polypeptides of the invention, and ability to bind to a receptor or ligand for a polypeptide of the invention.

5           The functional activity of polypeptides of the invention, and fragments, variants derivatives, and analogs thereof, can be assayed by various methods.

For example, in one embodiment where one is assaying for the ability to bind or compete with full-length polypeptide of the invention for binding to an antibody of the polypeptide of the invention, various immunoassays known in the art can be used, including but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, in situ immunoassays (using colloidal gold, enzyme or radioisotope labels, for example), western blots, precipitation reactions, agglutination assays (e.g., gel agglutination assays, hemagglutination assays), complement fixation assays, immunofluorescence assays, protein A assays, and immunoelectrophoresis assays, etc. In one embodiment, antibody binding is detected by detecting a label on the primary antibody. In another embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In a further embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are within the scope of the present invention.

10  
15  
20

In another embodiment, where a ligand for a polypeptide of the invention identified, or the ability of a polypeptide fragment, variant or derivative of the invention to multimerize is being evaluated, binding can be assayed, e.g., by means well-known in the art, such as, for example, reducing and non-reducing gel

25

chromatography, protein affinity chromatography, and affinity blotting. See generally, Phizicky, E., et al., 1995, Microbiol. Rev. 59:94-123. In another embodiment, physiological correlates of binding of a polypeptide of the invention to its substrates (signal transduction) can be assayed.

5           In addition, assays described herein (see Examples) and otherwise known in the art may routinely be applied to measure the ability of polypeptides of the invention and fragments, variants derivatives and analogs thereof to elicit related biological activity related to that of the polypeptide of the invention (either in vitro or in vivo). Other methods will be known to the skilled artisan and are within the scope  
10   of the invention.

### **Epitopes and Antibodies**

The present invention encompasses polypeptides comprising, or alternatively consisting of, an epitope of the polypeptide having an amino acid sequence of SEQ ID  
15   NO:Y, or an epitope of the polypeptide sequence encoded by a polynucleotide sequence contained in ATCC deposit No. Z or encoded by a polynucleotide that hybridizes to the complement of the sequence of SEQ ID NO:X or contained in ATCC deposit No. Z under stringent hybridization conditions or lower stringency hybridization conditions as defined supra. The present invention further encompasses  
20   polynucleotide sequences encoding an epitope of a polypeptide sequence of the invention (such as, for example, the sequence disclosed in SEQ ID NO:X), polynucleotide sequences of the complementary strand of a polynucleotide sequence encoding an epitope of the invention, and polynucleotide sequences which hybridize to the complementary strand under stringent hybridization conditions or lower  
25   stringency hybridization conditions defined supra.

The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in an animal, preferably a mammal, and most preferably in a human. In a preferred embodiment, the present invention encompasses a polypeptide comprising an epitope, as well as the polynucleotide encoding this polypeptide. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an antibody response in an animal, as determined by any method known in the art, for example, by the methods for generating antibodies described *infra*. (See, for example, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983)). The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody can immunospecifically bind its antigen as determined by any method well known in the art, for example, by the immunoassays described herein. Immunospecific binding excludes non-specific binding but does not necessarily exclude cross-reactivity with other antigens. Antigenic epitopes need not necessarily be immunogenic.

Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985), further described in U.S. Patent No. 4,631,211).

In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, and, most preferably, between about 15 to about 30 amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length. Additional non-exclusive preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as portions thereof. Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal antibodies,

that specifically bind the epitope. Preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these antigenic epitopes. Antigenic epitopes can be used as the target molecules in immunoassays. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe et al., Science 219:660-666 (1983)).

Similarly, immunogenic epitopes can be used, for example, to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle et al., J. Gen. Virol. 66:2347-2354 (1985). Preferred immunogenic epitopes include the immunogenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these immunogenic epitopes. The polypeptides comprising one or more immunogenic epitopes may be presented for eliciting an antibody response together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse), or, if the polypeptide is of sufficient length (at least about 25 amino acids), the polypeptide may be presented without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting).

Epitope-bearing polypeptides of the present invention may be used to induce antibodies according to methods well known in the art including, but not limited to, in vivo immunization, in vitro immunization, and phage display methods. See, e.g., Sutcliffe et al., supra; Wilson et al., supra, and Bittle et al., J. Gen. Virol., 66:2347-2354 (1985). If in vivo immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a

carrier using a linker such as maleimidobenzoyl- N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier- coupled peptides, for instance, by intraperitoneal and/or

5 intradermal injection of emulsions containing about 100  $\mu$ g of peptide or carrier protein and Freund's adjuvant or any other adjuvant known for stimulating an immune response. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid

10 surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

As one of skill in the art will appreciate, and as discussed above, the

15 polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to other polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins

20 may facilitate purification and may increase half-life in vivo. This has been shown for chimeric proteins consisting of the first two domains of the human CD4- polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., Nature, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the

25 immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fc fragments (see, e.g., PCT Publications WO

96/22024 and WO 99/04813). IgG Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion desulfide bonds have also been found to be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, e.g., Fountoulakis et al., J. Biochem., 5 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, 10 Proc. Natl. Acad. Sci. USA 88:8972- 897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with the recombinant vaccinia virus are loaded onto Ni<sup>2+</sup> 15 nitriloacetic acid-agarose column and histidine-tagged proteins can be selectively eluted with imidazole-containing buffers.

Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to 20 modulate the activities of polypeptides of the invention, such methods can be used to generate polypeptides with altered activity, as well as agonists and antagonists of the polypeptides. See, generally, U.S. Patent Nos. 5,605,793; 5,811,238; 5,830,721; 5,834,252; and 5,837,458, and Patten et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, Trends Biotechnol. 16(2):76-82 (1998); Hansson, et al., J. Mol. 25 Biol. 287:265-76 (1999); and Lorenzo and Blasco, Biotechniques 24(2):308- 13 (1998) (each of these patents and publications are hereby incorporated by reference in

its entirety). In one embodiment, alteration of polynucleotides corresponding to SEQ ID NO:X and the polypeptides encoded by these polynucleotides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments by homologous or site-specific recombination to generate variation in the polynucleotide sequence. In another embodiment, polynucleotides of the invention, or the encoded polypeptides, may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of a polynucleotide encoding a polypeptide of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules.

### Antibodies

Further polypeptides of the invention relate to antibodies and T-cell antigen receptors (TCR) which immunospecifically bind a polypeptide, polypeptide fragment, or variant of SEQ ID NO:Y, and/or an epitope, of the present invention (as determined by immunoassays well known in the art for assaying specific antibody-antigen binding). Antibodies of the invention include, but are not limited to, polyclonal, monoclonal, multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies (including, e.g., anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that immunospecifically binds an antigen. The immunoglobulin molecules of the invention can be of any type (e.g., IgG, IgE, IgM,

IgD, IgA and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule.

Most preferably the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')<sub>2</sub>, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, and CH3 domains. The antibodies of the invention may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine (e.g., mouse and rat), donkey, sheep rabbit, goat, guinea pig, camel, horse, or chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries or from animals transgenic for one or more human immunoglobulin and that do not express endogenous immunoglobulins, as described infra and, for example in, U.S. Patent No. 5,939,598 by Kucherlapati et al.

The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for a heterologous epitope, such as a heterologous polypeptide or solid support material. See, e.g., PCT publications WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); U.S. Patent Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny et al., J. Immunol. 148:1547-1553 (1992).



Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which they recognize or specifically bind. The epitope(s) or polypeptide portion(s) may be specified as described herein, e.g., by N-terminal and C-terminal positions, by size in  
5 contiguous amino acid residues, or listed in the Tables and Figures. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the exclusion of the same.

Antibodies of the present invention may also be described or specified in  
10 terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or homolog of a polypeptide of the present invention are included. Antibodies that bind polypeptides with at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, and at least 50% identity (as calculated using methods known in the art and described  
15 herein) to a polypeptide of the present invention are also included in the present invention. In specific embodiments, antibodies of the present invention cross-react with murine, rat and/or rabbit homologs of human proteins and the corresponding epitopes thereof. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%,  
20 less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In a specific embodiment, the above-described cross-reactivity is with respect to any single specific antigenic or immunogenic polypeptide, or combination(s) of 2, 3, 4, 5, or more of the specific  
25 antigenic and/or immunogenic polypeptides disclosed herein. Further included in the present invention are antibodies which bind polypeptides encoded by polynucleotides

which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity to a polypeptide of the invention. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, or  $10^{-15}$  M.

The invention also provides antibodies that competitively inhibit binding of an antibody to an epitope of the invention as determined by any method known in the art for determining competitive binding, for example, the immunoassays described herein. In preferred embodiments, the antibody competitively inhibits binding to the epitope by at least 95%, at least 90%, at least 85 %, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50%.

Antibodies of the present invention may act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Preferably, antibodies of the present invention bind an antigenic epitope disclosed herein, or a portion thereof. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or its substrate by immunoprecipitation followed by western blot analysis (for example, as described

supra). In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

5           The invention also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex, and, preferably, do not specifically recognize the unbound receptor or the unbound ligand. Likewise, included in the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as  
10   well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the  
15   receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides of the invention disclosed herein. The above antibody agonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Patent No. 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen et al., Cancer Res.  
20   58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon et al., J. Immunol. 160(7):3170-3179 (1998); Prat et al., J. Cell. Sci. 111(Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997); Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al.,  
25   Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998);

Bartunek et al., Cytokine 8(1):14-20 (1996) (which are all incorporated by reference herein in their entireties).

Antibodies of the present invention may be used, for example, but not limited to, to purify, detect, and target the polypeptides of the present invention, including both in vitro and in vivo diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference herein in its entirety).

As discussed in more detail below, the antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, radionuclides, or toxins. See, e.g., PCT publications WO 92/08495; WO 91/14438; WO 89/12624; U.S. Patent No. 5,314,995; and EP 396,387.

The antibodies of the invention include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques,

including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

The antibodies of the present invention may be generated by any suitable method known in the art. Polyclonal antibodies to an antigen-of- interest can be produced by various procedures well known in the art. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art.

Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Monoclonal Antibodies and T-Cell Hybridomas 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone,

including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art and are discussed in detail in the Examples (e.g., Example 16). In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing such peptide. Once an immune response is detected, e.g., antibodies specific for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding a polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by immunizing mice with positive hybridoma clones.

Accordingly, the present invention provides methods of generating monoclonal antibodies as well as antibodies produced by the method comprising culturing a hybridoma cell secreting an antibody of the invention wherein, preferably, the hybridoma is generated by fusing splenocytes isolated from a mouse immunized with an antigen of the invention with myeloma cells and then screening the hybridomas resulting from the fusion for hybridoma clones that secrete an antibody able to bind a polypeptide of the invention.

Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and F(ab')<sub>2</sub> fragments of the invention may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments).

F(ab')<sub>2</sub> fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

For example, the antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies,

including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')<sub>2</sub> fragments can also be employed using  
5 methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., *BioTechniques* 12(6):864-869 (1992); and Sawai et al., *AJRI* 34:26-34 (1995); and Better et al., *Science* 240:1041-1043 (1988) (said references incorporated by reference in their entirety).

Examples of techniques which can be used to produce single-chain Fvs and  
10 antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston et al., *Methods in Enzymology* 203:46-88 (1991); Shu et al., *PNAS* 90:7995-7999 (1993); and Skerra et al., *Science* 240:1038-1040 (1988). For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it may be preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule  
15 in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, *Science* 229:1202 (1985); Oi et al., *BioTechniques* 4:214 (1986); Gillies et al., (1989) *J. Immunol.*  
20 *Methods* 125:191-202; U.S. Patent Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having one or more complementarity determining regions (CDRs) from the non-human species and a framework regions from a human immunoglobulin molecule.  
25 Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve,



antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (See, e.g.,  
5 Queen et al., U.S. Patent No. 5,585,089; Riechmann et al., Nature 332:323 (1988), which are incorporated herein by reference in their entireties.) Antibodies can be humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Patent Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP  
10 519,596; Padlan, Molecular Immunology 28(4/5):489-498 (1991); Studnicka et al., Protein Engineering 7(6):805-814 (1994); Roguska. et al., PNAS 91:969-973 (1994)), and chain shuffling (U.S. Patent No. 5,565,332).

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies can be made by a variety of methods  
15 known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Patent Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety.

20 Human antibodies can also be produced using transgenic mice which are incapable of expressing functional endogenous immunoglobulins, but which can express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous recombination into mouse embryonic stem cells. Alternatively, the  
25 human variable region, constant region, and diversity region may be introduced into mouse embryonic stem cells in addition to the human heavy and light chain genes.

The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified  
5 embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained from the  
10 immunized, transgenic mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human  
15 antibodies, see Lonberg and Huszar, *Int. Rev. Immunol.* 13:65-93 (1995). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Patent Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825;  
20 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

25 Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a

selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., Bio/technology 12:899-903 (1988)).

Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotypic antibodies that “mimic” polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, FASEB J. 7(5):437-444; (1989) and Nissinoff, J. Immunol. 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to a ligand can be used to generate anti-idiotypes that “mimic” the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

#### Polynucleotides Encoding Antibodies

The invention further provides polynucleotides comprising a nucleotide sequence encoding an antibody of the invention and fragments thereof. The invention also encompasses polynucleotides that hybridize under stringent or lower stringency hybridization conditions, e.g., as defined supra, to polynucleotides that encode an antibody, preferably, that specifically binds to a polypeptide of the invention, preferably, an antibody that binds to a polypeptide having the amino acid sequence of SEQ ID NO:Y.

The polynucleotides may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. For example, if the

nucleotide sequence of the antibody is known, a polynucleotide encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., *BioTechniques* 17:242 (1994)), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence  
5 encoding the antibody, annealing and ligating of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

Alternatively, a polynucleotide encoding an antibody may be generated from nucleic acid from a suitable source. If a clone containing a nucleic acid encoding a particular antibody is not available, but the sequence of the antibody molecule is  
10 known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from, any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody of the invention) by PCR amplification using synthetic primers hybridizable  
15 to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, e.g., a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

20 Once the nucleotide sequence and corresponding amino acid sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook et al., 1990, *Molecular Cloning, A*  
25 *Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY and Ausubel et al., eds., 1998, *Current Protocols in Molecular Biology*, John

Wiley & Sons, NY, which are both incorporated by reference herein in their entireties ), to generate antibodies having a different amino acid sequence, for example to create amino acid substitutions, deletions, and/or insertions.

In a specific embodiment, the amino acid sequence of the heavy and/or light chain variable domains may be inspected to identify the sequences of the complementarity determining regions (CDRs) by methods that are well known in the art, e.g., by comparison to known amino acid sequences of other heavy and light chain variable regions to determine the regions of sequence hypervariability. Using routine recombinant DNA techniques, one or more of the CDRs may be inserted within framework regions, e.g., into human framework regions to humanize a non-human antibody, as described supra. The framework regions may be naturally occurring or consensus framework regions, and preferably human framework regions (see, e.g., Chothia et al., J. Mol. Biol. 278: 457-479 (1998) for a listing of human framework regions). Preferably, the polynucleotide generated by the combination of the framework regions and CDRs encodes an antibody that specifically binds a polypeptide of the invention. Preferably, as discussed supra, one or more amino acid substitutions may be made within the framework regions, and, preferably, the amino acid substitutions improve binding of the antibody to its antigen. Additionally, such methods may be used to make amino acid substitutions or deletions of one or more variable region cysteine residues participating in an intrachain disulfide bond to generate antibody molecules lacking one or more intrachain disulfide bonds. Other alterations to the polynucleotide are encompassed by the present invention and within the skill of the art.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., Proc. Natl. Acad. Sci. 81:851-855 (1984); Neuberger et al., Nature 312:604-608 (1984); Takeda et al., Nature 314:452-454 (1985)) by splicing genes

from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived  
5 from a murine mAb and a human immunoglobulin constant region, e.g., humanized antibodies.

Alternatively, techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778; Bird, Science 242:423- 42 (1988); Huston et al., Proc. Natl. Acad. Sci. USA 85:5879-5883 (1988); and Ward et al., Nature  
10 334:544-54 (1989)) can be adapted to produce single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in E. coli may also be used (Skerra et al., Science 242:1038- 1041 (1988)).

15

#### Methods of Producing Antibodies

The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

20 Recombinant expression of an antibody of the invention, or fragment, derivative or analog thereof, (e.g., a heavy or light chain of an antibody of the invention or a single chain antibody of the invention), requires construction of an expression vector containing a polynucleotide that encodes the antibody. Once a polynucleotide encoding an antibody molecule or a heavy or light chain of an  
25 antibody, or portion thereof (preferably containing the heavy or light chain variable domain), of the invention has been obtained, the vector for the production of the

antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing a protein by expressing a polynucleotide containing an antibody encoding nucleotide sequence are described herein. Methods which are well known to those skilled in the art can be  
5 used to construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. The invention, thus, provides replicable vectors comprising a nucleotide sequence encoding an antibody molecule of the invention, or a heavy or  
10 light chain thereof, or a heavy or light chain variable domain, operably linked to a promoter. Such vectors may include the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Patent No. 5,122,464) and the variable domain of the antibody may be cloned into such a vector for expression of the entire heavy or  
15 light chain.

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention. Thus, the invention includes host cells containing a polynucleotide encoding an antibody of the invention, or a heavy or light chain  
20 thereof, or a single chain antibody of the invention, operably linked to a heterologous promoter. In preferred embodiments for the expression of double-chained antibodies, vectors encoding both the heavy and light chains may be co-expressed in the host cell for expression of the entire immunoglobulin molecule, as detailed below.

A variety of host-expression vector systems may be utilized to express the  
25 antibody molecules of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently

purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express an antibody molecule of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO; BHK, 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Preferably, bacterial cells such as *Escherichia coli*, and more preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule, are used for the expression of a recombinant antibody molecule. For example, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., *Gene* 45:101 (1986); Cockett et al., *Bio/Technology* 8:2 (1990)).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the



generation of pharmaceutical compositions of an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., EMBO J. 2:1791 (1983)), in which the antibody coding  
5 sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, Nucleic Acids Res. 13:3101-3109 (1985); Van Heeke & Schuster, J. Biol. Chem. 24:5503-5509 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such  
10 fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

15 In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin  
20 promoter).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the antibody coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric  
25 gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region

E1 or E3) will result in a recombinant virus that is viable and capable of expressing the antibody molecule in infected hosts. (e.g., see Logan & Shenk, Proc. Natl. Acad. Sci. USA 81:355-359 (1984)). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the  
5 ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate  
10 transcription enhancer elements, transcription terminators, etc. (see Bittner et al., Methods in Enzymol. 153:51-544 (1987)).

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g.,  
15 cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which  
20 possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, and in particular, breast cancer cell lines such as, for example, BT483, Hs578T, HTB2, BT20 and T47D, and normal mammary gland cell  
25 line such as, for example, CRL7030 and Hs578Bst.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the antibody molecule may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the antibody molecule. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., Cell 11:223 (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA 48:202 (1992)), and adenine phosphoribosyltransferase (Lowy et al., Cell 22:817 (1980)) genes can be employed in tk-, hgp<sup>rt</sup>- or ap<sup>rt</sup>- cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., Natl. Acad. Sci. USA 77:357 (1980); O'Hare et al., Proc. Natl. Acad. Sci. USA 78:1527 (1981)); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci. USA 78:2072 (1981)); neo, which confers resistance to the aminoglycoside G-418 Clinical Pharmacy 12:488-505; Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science

260:926-932 (1993); and Morgan and Anderson, *Ann. Rev. Biochem.* 62:191-217 (1993); May, 1993, *TIB TECH* 11(5):155-215); and hygromycin, which confers resistance to hygromycin (Santerre et al., *Gene* 30:147 (1984)). Methods commonly known in the art of recombinant DNA technology may be routinely applied to select the desired recombinant clone, and such methods are described, for example, in Ausubel et al. (eds.), *Current Protocols in Molecular Biology*, John Wiley & Sons, NY (1993); Kriegler, *Gene Transfer and Expression, A Laboratory Manual*, Stockton Press, NY (1990); and in Chapters 12 and 13, Dracopoli et al. (eds), *Current Protocols in Human Genetics*, John Wiley & Sons, NY (1994); Colberre-Garapin et al., *J. Mol. Biol.* 150:1 (1981), which are incorporated by reference herein in their entireties.

The expression levels of an antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, *The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning*, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., *Mol. Cell. Biol.* 3:257 (1983)).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, *Nature* 322:52 (1986); Kohler, *Proc. Natl. Acad. Sci.*

USA 77:2197 (1980)). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

Once an antibody molecule of the invention has been produced by an animal, chemically synthesized, or recombinantly expressed, it may be purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (e.g., ion exchange, affinity, particularly by affinity for the specific antigen after Protein A, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. In addition, the antibodies of the present invention or fragments thereof can be fused to heterologous polypeptide sequences described herein or otherwise known in the art, to facilitate purification.

The present invention encompasses antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention to generate fusion proteins. The fusion does not necessarily need to be direct, but may occur through linker sequences. The antibodies may be specific for antigens other than polypeptides (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention. For example, antibodies may be used to target the polypeptides of the present invention to particular cell types, either in vitro or in vivo, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in in vitro immunoassays and purification methods using methods known in the art. See e.g., Harbor et al., supra, and PCT publication WO 93/21232; EP 439,095; Naramura et al., Immunol. Lett. 39:91-99 (1994); U.S. Patent 5,474,981; Gillies et al., PNAS

89:1428-1432 (1992); Fell et al., J. Immunol. 146:2446-2452(1991), which are incorporated by reference in their entireties.

The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the constant region, hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains or portions thereof. The polypeptides may also be fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. See, e.g., U.S. Patent Nos. 5,336,603; 5,622,929; 5,359,046; 5,349,053; 5,447,851; 5,112,946; EP 307,434; EP 367,166; PCT publications WO 96/04388; WO 91/06570; Ashkenazi et al., Proc. Natl. Acad. Sci. USA 88:10535-10539 (1991); Zheng et al., J. Immunol. 154:5590-5600 (1995); and Vil et al., Proc. Natl. Acad. Sci. USA 89:11337- 11341(1992) (said references incorporated by reference in their entireties).

As discussed, supra, the polypeptides corresponding to a polypeptide, polypeptide fragment, or a variant of SEQ ID NO:Y may be fused or conjugated to the above antibody portions to increase the in vivo half life of the polypeptides or for use in immunoassays using methods known in the art. Further, the polypeptides corresponding to SEQ ID NO:Y may be fused or conjugated to the above antibody portions to facilitate purification. One reported example describes chimeric proteins

consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP 394,827; Traunecker et al., *Nature* 331:84-86 (1988)). The polypeptides of the present invention fused or conjugated to an antibody having

5 disulfide- linked dimeric structures (due to the IgG) may also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995)). In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP A

10 232,262). Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to

15 identify antagonists of hIL-5. (See, Bennett et al., *J. Molecular Recognition* 8:52-58 (1995); Johanson et al., *J. Biol. Chem.* 270:9459-9471 (1995)).

Moreover, the antibodies or fragments thereof of the present invention can be fused to marker sequences, such as a peptide to facilitate purification. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the

20 tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Other peptide tags useful for purification include, but are not limited to, the "HA" tag, which

25 corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson et al., *Cell* 37:767 (1984)) and the "flag" tag.

The present invention further encompasses antibodies or fragments thereof conjugated to a diagnostic or therapeutic agent. The antibodies can be used diagnostically to, for example, monitor the development or progression of a tumor as part of a clinical testing procedure to, e.g., determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, and nonradioactive paramagnetic metal ions. The detectable substance may be coupled or conjugated either directly to the antibody (or fragment thereof) or indirectly, through an intermediate (such as, for example, a linker known in the art) using techniques known in the art. See, for example, U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{111}\text{In}$  or  $^{99}\text{Tc}$ .

Further, an antibody or fragment thereof may be conjugated to a therapeutic moiety such as a cytotoxin, e.g., a cytostatic or cytocidal agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example,  $^{213}\text{Bi}$ . A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include



paclitaxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and  
5 puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclophosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-  
10 dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological  
15 response, the therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor,  $\alpha$ -interferon,  $\beta$ -interferon, nerve growth  
20 factor, platelet derived growth factor, tissue plasminogen activator, an apoptotic agent, e.g., TNF- $\alpha$ , TNF- $\beta$ , AIM I (See, International Publication No. WO 97/33899), AIM II (See, International Publication No. WO 97/34911), Fas Ligand (Takahashi *et al.*, *Int. Immunol.*, 6:1567-1574 (1994)), VEGI (See, International Publication No. WO 99/23105), a thrombotic agent or an anti-angiogenic agent, e.g.,  
25 angiostatin or endostatin; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"),

granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Antibodies may also be attached to solid supports, which are particularly useful for immunoassays or purification of the target antigen. Such solid supports  
5 include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in *Monoclonal Antibodies And Cancer Therapy*, Reisfeld  
10 et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in *Controlled Drug Delivery* (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in *Monoclonal Antibodies '84: Biological And Clinical Applications*, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results,  
15 And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in *Monoclonal Antibodies For Cancer Detection And Therapy*, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", *Immunol. Rev.* 62:119-58 (1982).

20 Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980, which is incorporated herein by reference in its entirety.

An antibody, with or without a therapeutic moiety conjugated to it, administered alone or in combination with cytotoxic factor(s) and/or cytokine(s) can  
25 be used as a therapeutic.

### Immunophenotyping

The antibodies of the invention may be utilized for immunophenotyping of cell lines and biological samples. The translation product of the gene of the present invention may be useful as a cell specific marker, or more specifically as a cellular  
5 marker that is differentially expressed at various stages of differentiation and/or maturation of particular cell types. Monoclonal antibodies directed against a specific epitope, or combination of epitopes, will allow for the screening of cellular populations expressing the marker. Various techniques can be utilized using monoclonal antibodies to screen for cellular populations expressing the marker(s), and  
10 include magnetic separation using antibody-coated magnetic beads, "panning" with antibody attached to a solid matrix (i.e., plate), and flow cytometry (See, e.g., U.S. Patent 5,985,660; and Morrison *et al.*, *Cell*, 96:737-49 (1999)).

These techniques allow for the screening of particular populations of cells, such as might be found with hematological malignancies (i.e. minimal residual  
15 disease (MRD) in acute leukemic patients) and "non-self" cells in transplantations to prevent Graft-versus-Host Disease (GVHD). Alternatively, these techniques allow for the screening of hematopoietic stem and progenitor cells capable of undergoing proliferation and/or differentiation, as might be found in human umbilical cord blood.

### 20 Assays For Antibody Binding

The antibodies of the invention may be assayed for immunospecific binding by any method known in the art. The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent  
25 assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays,

complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few. Such assays are routine and well known in the art (see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York, which is incorporated by  
5 reference herein in its entirety). Exemplary immunoassays are described briefly below (but are not intended by way of limitation).

Immunoprecipitation protocols generally comprise lysing a population of cells in a lysis buffer such as RIPA buffer (1% NP-40 or Triton X- 100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 0.01 M sodium phosphate at pH 7.2, 1%  
10 Trasylol) supplemented with protein phosphatase and/or protease inhibitors (e.g., EDTA, PMSF, aprotinin, sodium vanadate), adding the antibody of interest to the cell lysate, incubating for a period of time (e.g., 1-4 hours) at 4° C, adding protein A and/or protein G sepharose beads to the cell lysate, incubating for about an hour or more at 4° C, washing the beads in lysis buffer and resuspending the beads in  
15 SDS/sample buffer. The ability of the antibody of interest to immunoprecipitate a particular antigen can be assessed by, e.g., western blot analysis. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the binding of the antibody to an antigen and decrease the background (e.g., pre-clearing the cell lysate with sepharose beads). For further discussion regarding  
20 immunoprecipitation protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.16.1.

Western blot analysis generally comprises preparing protein samples, electrophoresis of the protein samples in a polyacrylamide gel (e.g., 8%- 20% SDS-PAGE depending on the molecular weight of the antigen), transferring the protein  
25 sample from the polyacrylamide gel to a membrane such as nitrocellulose, PVDF or nylon, blocking the membrane in blocking solution (e.g., PBS with 3% BSA or non-

fat milk), washing the membrane in washing buffer (e.g., PBS-Tween 20), blocking the membrane with primary antibody (the antibody of interest) diluted in blocking buffer, washing the membrane in washing buffer, blocking the membrane with a secondary antibody (which recognizes the primary antibody, e.g., an anti-human  
5 antibody) conjugated to an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) or radioactive molecule (e.g.,  $^{32}\text{P}$  or  $^{125}\text{I}$ ) diluted in blocking buffer, washing the membrane in wash buffer, and detecting the presence of the antigen. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected and to reduce the background noise. For  
10 further discussion regarding western blot protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.8.1.

ELISAs comprise preparing antigen, coating the well of a 96 well microtiter plate with the antigen, adding the antibody of interest conjugated to a detectable  
15 compound such as an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) to the well and incubating for a period of time, and detecting the presence of the antigen. In ELISAs the antibody of interest does not have to be conjugated to a detectable compound; instead, a second antibody (which recognizes the antibody of interest) conjugated to a detectable compound may be added to the  
20 well. Further, instead of coating the well with the antigen, the antibody may be coated to the well. In this case, a second antibody conjugated to a detectable compound may be added following the addition of the antigen of interest to the coated well. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected as well as other variations of ELISAs  
25 known in the art. For further discussion regarding ELISAs see, e.g., Ausubel et al,

eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 11.2.1.

The binding affinity of an antibody to an antigen and the off-rate of an antibody-antigen interaction can be determined by competitive binding assays. One  
5 example of a competitive binding assay is a radioimmunoassay comprising the incubation of labeled antigen (e.g., 3H or 125I) with the antibody of interest in the presence of increasing amounts of unlabeled antigen, and the detection of the antibody bound to the labeled antigen. The affinity of the antibody of interest for a particular antigen and the binding off-rates can be determined from the data by  
10 scatchard plot analysis. Competition with a second antibody can also be determined using radioimmunoassays. In this case, the antigen is incubated with antibody of interest conjugated to a labeled compound (e.g., 3H or 125I) in the presence of increasing amounts of an unlabeled second antibody.

### 15        Therapeutic Uses

The present invention is further directed to antibody-based therapies which involve administering antibodies of the invention to an animal, preferably a mammal, and most preferably a human, patient for treating one or more of the disclosed diseases, disorders, or conditions. Therapeutic compounds of the invention include,  
20 but are not limited to, antibodies of the invention (including fragments, analogs and derivatives thereof as described herein) and nucleic acids encoding antibodies of the invention (including fragments, analogs and derivatives thereof and anti-idiotypic antibodies as described herein). The antibodies of the invention can be used to treat, inhibit or prevent diseases, disorders or conditions associated with aberrant expression  
25 and/or activity of a polypeptide of the invention, including, but not limited to, any one or more of the diseases, disorders, or conditions described herein. The treatment

and/or prevention of diseases, disorders, or conditions associated with aberrant expression and/or activity of a polypeptide of the invention includes, but is not limited to, alleviating symptoms associated with those diseases, disorders or conditions. Antibodies of the invention may be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors (such as, e.g., IL-2, IL-3 and IL-7), for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

The antibodies of the invention may be administered alone or in combination with other types of treatments (e.g., radiation therapy, chemotherapy, hormonal therapy, immunotherapy and anti-tumor agents). Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, human antibodies, fragments derivatives, analogs, or nucleic acids, are administered to a human patient for therapy or prophylaxis.

It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of disorders related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides of the invention, including fragments thereof. Preferred binding affinities include those with a dissociation constant or  $K_d$  less than  $5 \times 10^{-2}$  M,  $10^{-2}$  M,  $5 \times 10^{-3}$  M,  $10^{-3}$  M,  $5 \times 10^{-4}$  M,  $10^{-4}$  M,  $5 \times 10^{-5}$  M,  $10^{-5}$  M,  $5 \times 10^{-6}$  M,  $10^{-6}$  M,  $5 \times 10^{-7}$  M,  $10^{-7}$  M,  $5 \times 10^{-8}$  M,  $10^{-8}$  M,  $5 \times 10^{-9}$  M,  $10^{-9}$  M,  $5 \times 10^{-10}$  M,  $10^{-10}$  M,  $5 \times 10^{-11}$  M,  $10^{-11}$  M,  $5 \times 10^{-12}$  M,  $10^{-12}$  M,  $5 \times 10^{-13}$  M,  $10^{-13}$  M,  $5 \times 10^{-14}$  M,  $10^{-14}$  M,  $5 \times 10^{-15}$  M, and  $10^{-15}$  M.

### Gene Therapy

In a specific embodiment, nucleic acids comprising sequences encoding antibodies or functional derivatives thereof, are administered to treat, inhibit or prevent a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention, by way of gene therapy. Gene therapy refers to therapy performed by the administration to a subject of an expressed or expressible nucleic acid. In this embodiment of the invention, the nucleic acids produce their encoded protein that mediates a therapeutic effect.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

For general reviews of the methods of gene therapy, see Goldspiel et al., Clinical Pharmacy 12:488-505 (1993); Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science 260:926-932 (1993); and Morgan and Anderson, Ann. Rev. Biochem. 62:191-217



(1993); May, TIBTECH 11(5):155-215 (1993). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, NY (1993); and Kriegler, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY  
5 (1990).

In a preferred aspect, the compound comprises nucleic acid sequences encoding an antibody, said nucleic acid sequences being part of expression vectors that express the antibody or fragments or chimeric proteins or heavy or light chains thereof in a suitable host. In particular, such nucleic acid sequences have promoters  
10 operably linked to the antibody coding region, said promoter being inducible or constitutive, and, optionally, tissue- specific. In another particular embodiment, nucleic acid molecules are used in which the antibody coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the  
15 antibody encoding nucleic acids (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al., Nature 342:435-438 (1989). In specific embodiments, the expressed antibody molecule is a single chain antibody; alternatively, the nucleic acid sequences include sequences encoding both the heavy and light chains, or fragments thereof, of the antibody.

20 Delivery of the nucleic acids into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid- carrying vectors, or indirect, in which case, cells are first transformed with the nucleic acids in vitro, then transplanted into the patient. These two approaches are known, respectively, as in vivo or ex vivo gene therapy.

25 In a specific embodiment, the nucleic acid sequences are directly administered in vivo, where it is expressed to produce the encoded product. This can be

accomplished by any of numerous methods known in the art, e.g., by constructing them as part of an appropriate nucleic acid expression vector and administering it so that they become intracellular, e.g., by infection using defective or attenuated retrovirals or other viral vectors (see U.S. Patent No. 4,980,286), or by direct  
5 injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering them in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see,  
10 e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, nucleic acid-ligand complexes can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo for  
15 cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188, WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al.,  
20 Nature 342:435-438 (1989)).

In a specific embodiment, viral vectors that contains nucleic acid sequences encoding an antibody of the invention are used. For example, a retroviral vector can be used (see Miller et al., Meth. Enzymol. 217:581-599 (1993)). These retroviral vectors contain the components necessary for the correct packaging of the viral  
25 genome and integration into the host cell DNA. The nucleic acid sequences encoding the antibody to be used in gene therapy are cloned into one or more vectors, which

facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., *Biotherapy* 6:291-302 (1994), which describes the use of a retroviral vector to deliver the *mdr1* gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., *J. Clin. Invest.* 93:644-651 (1994); Kiem et al., *Blood* 83:1467-1473 (1994); Salmons and Gunzberg, *Human Gene Therapy* 4:129-141 (1993); and Grossman and Wilson, *Curr. Opin. in Genetics and Devel.* 3:110-114 (1993).

Adenoviruses are other viral vectors that can be used in gene therapy.

Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, *Current Opinion in Genetics and Development* 3:499-503 (1993) present a review of adenovirus-based gene therapy. Bout et al., *Human Gene Therapy* 5:3-10 (1994) demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., *Science* 252:431-434 (1991); Rosenfeld et al., *Cell* 68:143-155 (1992); Mastrangeli et al., *J. Clin. Invest.* 91:225-234 (1993); PCT Publication WO94/12649; and Wang, et al., *Gene Therapy* 2:775-783 (1995). In a preferred embodiment, adenovirus vectors are used.

Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., *Proc. Soc. Exp. Biol. Med.* 204:289-300 (1993); U.S. Patent No. 5,436,146).

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection  
5 to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection,  
10 electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, Meth. Enzymol. 217:599-618 (1993); Cohen et al., Meth.  
15 Enzymol. 217:618-644 (1993); Cline, Pharmac. Ther. 29:69-92m (1985) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and  
20 expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be  
25 determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as Tlymphocytes, Blymphocytes, monocytes, macrophages, 5 neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

10 In an embodiment in which recombinant cells are used in gene therapy, nucleic acid sequences encoding an antibody are introduced into the cells such that they are expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and 15 maintained in vitro can potentially be used in accordance with this embodiment of the present invention (see e.g. PCT Publication WO 94/08598; Stemple and Anderson, Cell 71:973-985 (1992); Rheinwald, Meth. Cell Bio. 21A:229 (1980); and Pittelkow and Scott, Mayo Clinic Proc. 61:771 (1986)).

In a specific embodiment, the nucleic acid to be introduced for purposes of 20 gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription. Demonstration of Therapeutic or Prophylactic Activity

The compounds or pharmaceutical compositions of the invention are 25 preferably tested in vitro, and then in vivo for the desired therapeutic or prophylactic activity, prior to use in humans. For example, in vitro assays to demonstrate the

therapeutic or prophylactic utility of a compound or pharmaceutical composition include, the effect of a compound on a cell line or a patient tissue sample. The effect of the compound or composition on the cell line and/or tissue sample can be determined utilizing techniques known to those of skill in the art including, but not limited to, rosette formation assays and cell lysis assays. In accordance with the invention, in vitro assays which can be used to determine whether administration of a specific compound is indicated, include in vitro cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a compound, and the effect of such compound upon the tissue sample is observed.

10

#### Therapeutic/Prophylactic Administration and Composition

The invention provides methods of treatment, inhibition and prophylaxis by administration to a subject of an effective amount of a compound or pharmaceutical composition of the invention, preferably an antibody of the invention. In a preferred aspect, the compound is substantially purified (e.g., substantially free from substances that limit its effect or produce undesired side-effects). The subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human.

15

Formulations and methods of administration that can be employed when the compound comprises a nucleic acid or an immunoglobulin are described above; additional appropriate formulations and routes of administration can be selected from among those described herein below.

20

Various delivery systems are known and can be used to administer a compound of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (see, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)),

25

construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds or compositions may be administered by any convenient route, for example by infusion  
5 or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compounds or compositions of the invention into the central nervous system by any suitable route,  
10 including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

In a specific embodiment, it may be desirable to administer the pharmaceutical  
15 compounds or compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material,  
20 including membranes, such as sialastic membranes, or fibers. Preferably, when administering a protein, including an antibody, of the invention, care must be taken to use materials to which the protein does not absorb.

In another embodiment, the compound or composition can be delivered in a vesicle, in particular a liposome (see Langer, Science 249:1527-1533 (1990); Treat et  
25 al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein

and Fidler (eds.), Liss, New York, pp. 353- 365 (1989); Lopez-Berestein, *ibid.*, pp. 317-327; see generally *ibid.*)

In yet another embodiment, the compound or composition can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, 5 *supra*; Sefton, *CRC Crit. Ref. Biomed. Eng.* 14:201 (1987); Buchwald et al., *Surgery* 88:507 (1980); Saudek et al., *N. Engl. J. Med.* 321:574 (1989)). In another embodiment, polymeric materials can be used (see *Medical Applications of Controlled Release*, Langer and Wise (eds.), CRC Pres., Boca Raton, Florida (1974); *Controlled Drug Bioavailability, Drug Product Design and Performance*, Smolen and 10 Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., *Macromol. Sci. Rev. Macromol. Chem.* 23:61 (1983); see also Levy et al., *Science* 228:190 (1985); During et al., *Ann. Neurol.* 25:351 (1989); Howard et al., *J. Neurosurg.* 71:105 (1989)). In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose 15 (see, e.g., Goodson, in *Medical Applications of Controlled Release*, *supra*, vol. 2, pp. 115-138 (1984)).

Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)).

In a specific embodiment where the compound of the invention is a nucleic 20 acid encoding a protein, the nucleic acid can be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Patent No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating 25 with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox- like peptide which is known to enter the nucleus (see e.g.,



Joliot et al., Proc. Natl. Acad. Sci. USA 88:1864-1868 (1991)), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such  
5 compositions comprise a therapeutically effective amount of a compound, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term  
10 "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline  
15 solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired,  
20 can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as  
25 pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable

pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E.W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation  
5 should suit the mode of administration.

In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the  
10 composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the  
15 composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

The compounds of the invention can be formulated as neutral or salt forms.  
20 Pharmaceutically acceptable salts include those formed with anions such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

25 The amount of the compound of the invention which will be effective in the treatment, inhibition and prevention of a disease or disorder associated with aberrant

expression and/or activity of a polypeptide of the invention can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

For antibodies, the dosage administered to a patient is typically 0.1 mg/kg to 100 mg/kg of the patient's body weight. Preferably, the dosage administered to a patient is between 0.1 mg/kg and 20 mg/kg of the patient's body weight, more preferably 1 mg/kg to 10 mg/kg of the patient's body weight. Generally, human antibodies have a longer half-life within the human body than antibodies from other species due to the immune response to the foreign polypeptides. Thus, lower dosages of human antibodies and less frequent administration is often possible. Further, the dosage and frequency of administration of antibodies of the invention may be reduced by enhancing uptake and tissue penetration (e.g., into the brain) of the antibodies by modifications such as, for example, lipidation.

The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. **Diagnosis and Imaging**

Labeled antibodies, and derivatives and analogs thereof, which specifically bind to a polypeptide of interest can be used for diagnostic purposes to detect,

diagnose, or monitor diseases, disorders, and/or conditions associated with the aberrant expression and/or activity of a polypeptide of the invention. The invention provides for the detection of aberrant expression of a polypeptide of interest, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of aberrant expression.

The invention provides a diagnostic assay for diagnosing a disorder, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a particular disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Antibodies of the invention can be used to assay protein levels in a biological sample using classical immunohistological methods known to those of skill in the art (e.g., see Jalkanen, et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, et al., J. Cell Biol. 105:3087-3096 (1987)). Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody

assay labels are known in the art and include enzyme labels, such as, glucose oxidase; radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and technetium ( $^{99}\text{Tc}$ ); luminescent labels, such as luminol; and fluorescent labels, such as fluorescein and rhodamine, and biotin.

5           One aspect of the invention is the detection and diagnosis of a disease or disorder associated with aberrant expression of a polypeptide of interest in an animal, preferably a mammal and most preferably a human. In one embodiment, diagnosis comprises: a) administering (for example, parenterally, subcutaneously, or intraperitoneally) to a subject an effective amount of a labeled molecule which  
10           specifically binds to the polypeptide of interest; b) waiting for a time interval following the administering for permitting the labeled molecule to preferentially concentrate at sites in the subject where the polypeptide is expressed (and for unbound labeled molecule to be cleared to background level); c) determining background level; and d) detecting the labeled molecule in the subject, such that  
15           detection of labeled molecule above the background level indicates that the subject has a particular disease or disorder associated with aberrant expression of the polypeptide of interest. Background level can be determined by various methods including, comparing the amount of labeled molecule detected to a standard value previously determined for a particular system.

20           It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{mTc}}$ . The labeled antibody or antibody fragment will then preferentially  
25           accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of

Radiolabeled Antibodies and Their Fragments.” (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).

Depending on several variables, including the type of label used and the mode  
5 of administration, the time interval following the administration for permitting the labeled molecule to preferentially concentrate at sites in the subject and for unbound labeled molecule to be cleared to background level is 6 to 48 hours or 6 to 24 hours or 6 to 12 hours. In another embodiment the time interval following administration is 5 to 20 days or 5 to 10 days.

10 In an embodiment, monitoring of the disease or disorder is carried out by repeating the method for diagnosing the disease or disease, for example, one month after initial diagnosis, six months after initial diagnosis, one year after initial diagnosis, etc.

Presence of the labeled molecule can be detected in the patient using methods  
15 known in the art for in vivo scanning. These methods depend upon the type of label used. Skilled artisans will be able to determine the appropriate method for detecting a particular label. Methods and devices that may be used in the diagnostic methods of the invention include, but are not limited to, computed tomography (CT), whole body scan such as position emission tomography (PET), magnetic resonance imaging  
20 (MRI), and sonography.

In a specific embodiment, the molecule is labeled with a radioisotope and is detected in the patient using a radiation responsive surgical instrument (Thurston et al., U.S. Patent No. 5,441,050). In another embodiment, the molecule is labeled with a fluorescent compound and is detected in the patient using a fluorescence responsive  
25 scanning instrument. In another embodiment, the molecule is labeled with a positron emitting metal and is detected in the patent using positron emission-tomography. In

yet another embodiment, the molecule is labeled with a paramagnetic label and is detected in a patient using magnetic resonance imaging (MRI). Kits

The present invention provides kits that can be used in the above methods. In one embodiment, a kit comprises an antibody of the invention, preferably a purified  
5 antibody, in one or more containers. In a specific embodiment, the kits of the present invention contain a substantially isolated polypeptide comprising an epitope which is specifically immunoreactive with an antibody included in the kit. Preferably, the kits of the present invention further comprise a control antibody which does not react with the polypeptide of interest. In another specific embodiment, the kits of the present  
10 invention contain a means for detecting the binding of an antibody to a polypeptide of interest (e.g., the antibody may be conjugated to a detectable substrate such as a fluorescent compound, an enzymatic substrate, a radioactive compound or a luminescent compound, or a second antibody which recognizes the first antibody may be conjugated to a detectable substrate).

15 In another specific embodiment of the present invention, the kit is a diagnostic kit for use in screening serum containing antibodies specific against proliferative and/or cancerous polynucleotides and polypeptides. Such a kit may include a control antibody that does not react with the polypeptide of interest. Such a kit may include a substantially isolated polypeptide antigen comprising an epitope which is specifically  
20 immunoreactive with at least one anti-polypeptide antigen antibody. Further, such a kit includes means for detecting the binding of said antibody to the antigen (e.g., the antibody may be conjugated to a fluorescent compound such as fluorescein or rhodamine which can be detected by flow cytometry). In specific embodiments, the kit may include a recombinantly produced or chemically synthesized polypeptide  
25 antigen. The polypeptide antigen of the kit may also be attached to a solid support.

In a more specific embodiment the detecting means of the above-described kit includes a solid support to which said polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment, binding of the antibody to the polypeptide antigen can be detected by  
5 binding of the said reporter-labeled antibody.

In an additional embodiment, the invention includes a diagnostic kit for use in screening serum containing antigens of the polypeptide of the invention. The diagnostic kit includes a substantially isolated antibody specifically immunoreactive with polypeptide or polynucleotide antigens, and means for detecting the binding of  
10 the polynucleotide or polypeptide antigen to the antibody. In one embodiment, the antibody is attached to a solid support. In a specific embodiment, the antibody may be a monoclonal antibody. The detecting means of the kit may include a second, labeled monoclonal antibody. Alternatively, or in addition, the detecting means may include a labeled, competing antigen.

15 In one diagnostic configuration, test serum is reacted with a solid phase reagent having a surface-bound antigen obtained by the methods of the present invention. After binding with specific antigen antibody to the reagent and removing unbound serum components by washing, the reagent is reacted with reporter-labeled anti-human antibody to bind reporter to the reagent in proportion to the amount of  
20 bound anti-antigen antibody on the solid support. The reagent is again washed to remove unbound labeled antibody, and the amount of reporter associated with the reagent is determined. Typically, the reporter is an enzyme which is detected by incubating the solid phase in the presence of a suitable fluorometric, luminescent or colorimetric substrate (Sigma, St. Louis, MO).

25 The solid surface reagent in the above assay is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip



sticks, 96-well plate or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group.

- 5 Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

Thus, the invention provides an assay system or kit for carrying out this diagnostic method. The kit generally includes a support with surface-bound recombinant antigens, and a reporter-labeled anti-human antibody for detecting  
10 surface-bound anti-antigen antibody.

### **Fusion Proteins**

Any polypeptide of the present invention can be used to generate fusion  
15 proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because secreted proteins target cellular locations based on trafficking signals, the polypeptides of the present  
20 invention can be used as targeting molecules once fused to other proteins.

Examples of domains that can be fused to polypeptides of the present invention include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

25 Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino

acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final  
5 preparation of the polypeptide. The addition of peptide moieties to facilitate handling of polypeptides are familiar and routine techniques in the art.

Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgA, IgE, IgG, IgM) or portions thereof (CH1, CH2, CH3, and any  
10 combination thereof, including both entire domains and portions thereof), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian  
15 immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

20 Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively,  
25 deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if

the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

Moreover, the polypeptides of the present invention can be fused to marker sequences, such as a peptide which facilitates purification of the fused polypeptide. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984).)

Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

### **Vectors, Host Cells, and Protein Production**

The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line  
5 and then transduced into host cells.

The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The  
10 expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

15 As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*,  
20 *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

25 Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A,

pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlsbad, CA). Other suitable vectors will be readily apparent to the skilled artisan.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical

synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated  
5 or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal  
10 methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

In one embodiment, the yeast *Pichia pastoris* is used to express the polypeptide of the present invention in a eukaryotic system. *Pichia pastoris* is a  
15 methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolization pathway is the oxidation of methanol to formaldehyde using O<sub>2</sub>. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol  
20 oxidase for O<sub>2</sub>. Consequently, in a growth medium depending on methanol as a main carbon source, the promoter region of one of the two alcohol oxidase genes (*AOX1*) is highly active. In the presence of methanol, alcohol oxidase produced from the *AOX1* gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. See, Ellis, S.B., et al., *Mol. Cell. Biol.* 5:1111-21 (1985); Koutz, P.J., et al.,  
25 *Yeast* 5:167-77 (1989); Tschopp, J.F., et al., *Nucl. Acids Res.* 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a polynucleotide of the

present invention, under the transcriptional regulation of all or part of the *AOX1* regulatory sequence is expressed at exceptionally high levels in *Pichia* yeast grown in the presence of methanol.

In one example, the plasmid vector pPIC9K is used to express DNA encoding  
5 a polypeptide of the invention, as set forth herein; in a *Pichea* yeast system essentially as described in "*Pichia* Protocols: Methods in Molecular Biology," D.R. Higgins and J. Cregg, eds. The Humana Press, Totowa, NJ, 1998. This expression vector allows expression and secretion of a protein of the invention by virtue of the strong *AOX1* promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal  
10 peptide (i.e., leader) located upstream of a multiple cloning site.

Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily appreciate, as long as the proposed expression construct provides appropriately  
15 located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG as required.

In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, may be achieved by cloning the heterologous polynucleotide of the invention into an  
20 expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have  
25 been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide

sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination, resulting in the formation of a new transcription unit (see, e.g., U.S. Patent No. 5,641,670, issued June 24, 1997; U.S. Patent No. 5,733,761, issued March 31, 1998; International Publication No. WO 96/29411, published September 26, 1996; International Publication No. WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, *Proteins: Structures and Molecular Principles*, W.H. Freeman & Co., N.Y., and Hunkapiller et al., *Nature*, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide sequence of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid,  $\alpha$ -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid,  $\gamma$ -Abu,  $\epsilon$ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine,  $\beta$ -alanine, fluoro-amino acids, designer amino acids such as  $\beta$ -methyl amino acids,  $\alpha$ -methyl amino



acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

The invention encompasses polypeptides which are differentially modified during or after translation, *e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH<sub>4</sub>; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

Additional post-translational modifications encompassed by the invention include, for example, *e.g.*, N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of procaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein.

Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Patent NO: 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random

positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

The polymer may be of any molecular weight, and may be branched or unbranched. For polyethylene glycol, the preferred molecular weight is between  
5 about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or  
10 lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those  
15 skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated  
20 polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for  
25 therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules.

Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminal) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, *Therapeutics*) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of SEQ ID NO:Y or encoded by the cDNA contained in a deposited clone (including fragments, variants, splice variants, and

fusion proteins, corresponding to these polypeptides as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another  
5 specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (*e.g.*, containing polypeptides having identical or different amino acid sequences) or a homotrimer (*e.g.*, containing polypeptides having identical and/or different amino acid sequences). In additional  
10 embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (*i.e.*, polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the  
15 invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example,  
20 liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the  
25 invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the

invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence ( e.g., that recited in the sequence listing, or contained in the polypeptide encoded by a deposited clone). In one  
5 instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the  
10 heterologous polypeptide sequence in a fusion protein of the invention.

In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., US Patent Number 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in an Fc fusion protein of the invention (as  
15 described herein). In another specific example, covalent associations of fusion proteins of the invention are between heterologous polypeptide sequence from another protein that is capable of forming covalently associated multimers, such as for example, osteoprotegerin (see, e.g., International Publication NO: WO 98/49305, the contents of which are herein incorporated by reference in its entirety). In another  
20 embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627 (hereby incorporated by reference). Proteins comprising multiple polypeptides of the invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

25 Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper

polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using techniques known in the art.

Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide sequence. In a further embodiment, associations proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., US Patent  
5 Number 5,478,925, which is herein incorporated by reference in its entirety).  
Additionally, multimers of the invention may be generated using techniques known in the art to form one or more inter-molecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., US Patent Number 5,478,925, which is herein incorporated by  
10 reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). Additionally,  
15 techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

Alternatively, multimers of the invention may be generated using genetic  
20 engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are  
25 generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic

polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety). In another embodiment, recombinant techniques described  
5 herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., US Patent Number 5,478,925, which is herein incorporated by reference in its entirety).

10

#### **Uses of the Polynucleotides**

Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

15

The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

20

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the sequences shown in SEQ ID NO:X. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids  
25 containing the human gene corresponding to the SEQ ID NO:X will yield an amplified fragment.



Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. (Disease mapping data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library) .) Assuming 1 megabase mapping resolution and

one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected individuals can be examined.

5 First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected individuals, but not in normal individuals, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide  
10 and the corresponding gene from several normal individuals is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

Furthermore, increased or decreased expression of the gene in affected individuals as compared to unaffected individuals can be assessed using  
15 polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

Thus, the invention also provides a diagnostic method useful during diagnosis of a disorder, involving measuring the expression level of polynucleotides of the  
20 present invention in cells or body fluid from an individual and comparing the measured gene expression level with a standard level of polynucleotide expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a disorder.

In still another embodiment, the invention includes a kit for analyzing samples  
25 for the presence of proliferative and/or cancerous polynucleotides derived from a test subject. In a general embodiment, the kit includes at least one polynucleotide probe

containing a nucleotide sequence that will specifically hybridize with a polynucleotide of the present invention and a suitable container. In a specific embodiment, the kit includes two polynucleotide probes defining an internal region of the polynucleotide of the present invention, where each probe has one strand  
5 containing a 31' mer-end internal to the region. In a further embodiment, the probes may be useful as primers for polymerase chain reaction amplification.

Where a diagnosis of a disorder, has already been made according to conventional methods, the present invention is useful as a prognostic indicator, whereby patients exhibiting enhanced or depressed polynucleotide of the present  
10 invention expression will experience a worse clinical outcome relative to patients expressing the gene at a level nearer the standard level.

By "measuring the expression level of polynucleotide of the present invention" is intended qualitatively or quantitatively measuring or estimating the level of the polypeptide of the present invention or the level of the mRNA encoding the  
15 polypeptide in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the polypeptide level or mRNA level in a second biological sample). Preferably, the polypeptide level or mRNA level in the first biological sample is measured or estimated and compared to a standard polypeptide level or mRNA level, the standard  
20 being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of individuals not having a disorder. As will be appreciated in the art, once a standard polypeptide level or mRNA level is known, it can be used repeatedly as a standard for comparison.

25 By "biological sample" is intended any biological sample obtained from an individual, body fluid, cell line, tissue culture, or other source which contains the

polypeptide of the present invention or mRNA. As indicated, biological samples include body fluids (such as semen, lymph, sera, plasma, urine, synovial fluid and spinal fluid) which contain the polypeptide of the present invention, and other tissue sources found to express the polypeptide of the present invention. Methods for  
5 obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

The method(s) provided above may preferably be applied in a diagnostic method and/or kits in which polynucleotides and/or polypeptides are attached to a  
10 solid support. In one exemplary method, the support may be a "gene chip" or a "biological chip" as described in US Patents 5,837,832, 5,874,219, and 5,856,174. Further, such a gene chip with polynucleotides of the present invention attached may be used to identify polymorphisms between the polynucleotide sequences, with polynucleotides isolated from a test subject. The knowledge of such polymorphisms  
15 (i.e. their location, as well as, their existence) would be beneficial in identifying disease loci for many disorders, including cancerous diseases and conditions. Such a method is described in US Patents 5,858,659 and 5,856,104. The US Patents referenced supra are hereby incorporated by reference in their entirety herein.

The present invention encompasses polynucleotides of the present invention  
20 that are chemically synthesized, or reproduced as peptide nucleic acids (PNA), or according to other methods known in the art. The use of PNAs would serve as the preferred form if the polynucleotides are incorporated onto a solid support, or gene chip. For the purposes of the present invention, a peptide nucleic acid (PNA) is a polyamide type of DNA analog and the monomeric units for adenine, guanine,  
25 thymine and cytosine are available commercially (Perceptive Biosystems). Certain components of DNA, such as phosphorus, phosphorus oxides, or deoxyribose

derivatives, are not present in PNAs. As disclosed by P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, Science 254, 1497 (1991); and M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, Nature 365, 666 (1993), PNAs bind specifically and tightly to complementary DNA strands and are not degraded by nucleases. In fact, PNA binds more strongly to DNA than DNA itself does. This is probably because there is no electrostatic repulsion between the two strands, and also the polyamide backbone is more flexible. Because of this, PNA/DNA duplexes bind under a wider range of stringency conditions than DNA/DNA duplexes, making it easier to perform multiplex hybridization. Smaller probes can be used than with DNA due to the strong binding. In addition, it is more likely that single base mismatches can be determined with PNA/DNA hybridization because a single mismatch in a PNA/DNA 15-mer lowers the melting point ( $T_{sub.m}$ ) by 8°-20° C, vs. 4°-16° C for the DNA/DNA 15-mer duplex. Also, the absence of charge groups in PNA means that hybridization can be done at low ionic strengths and reduce possible interference by salt during the analysis.

The present invention is useful for detecting cancer in mammals. In particular the invention is useful during diagnosis of pathological cell proliferative neoplasias which include, but are not limited to: acute myelogenous leukemias including acute monocytic leukemia, acute myeloblastic leukemia, acute promyelocytic leukemia, acute myelomonocytic leukemia, acute erythroleukemia, acute megakaryocytic leukemia, and acute undifferentiated leukemia, etc.; and chronic myelogenous leukemias including chronic myelomonocytic leukemia, chronic granulocytic leukemia, etc. Preferred mammals include monkeys, apes, cats, dogs, cows, pigs, horses, rabbits and humans. Particularly preferred are humans.

Pathological cell proliferative diseases, disorders, and/or conditions are often associated with inappropriate activation of proto-oncogenes. (Germann, E. P. et al., "The Etiology of Acute Leukemia: Molecular Genetics and Viral Oncology," in Neoplastic Diseases of the Blood, Vol 1., Wiernik, P. H. et al. eds., 161-182 (1985)).

- 5 Neoplasias are now believed to result from the qualitative alteration of a normal cellular gene product, or from the quantitative modification of gene expression by insertion into the chromosome of a viral sequence, by chromosomal translocation of a gene to a more actively transcribed region, or by some other mechanism. (Germann et al., supra) It is likely that mutated or altered expression of specific genes is
- 10 involved in the pathogenesis of some leukemias, among other tissues and cell types. (Germann et al., supra) Indeed, the human counterparts of the oncogenes involved in some animal neoplasias have been amplified or translocated in some cases of human leukemia and carcinoma. (Germann et al., supra)
- For example, c-myc expression is highly amplified in the non-lymphocytic leukemia
- 15 cell line HL-60. When HL-60 cells are chemically induced to stop proliferation, the level of c-myc is found to be downregulated. (International Publication Number WO 91/15580) However, it has been shown that exposure of HL-60 cells to a DNA construct that is complementary to the 5' end of c-myc or c-myb blocks translation of the corresponding mRNAs which downregulates expression of the c-myc or c-myb
- 20 proteins and causes arrest of cell proliferation and differentiation of the treated cells. (International Publication Number WO 91/15580; Wickstrom et al., Proc. Natl. Acad. Sci. 85:1028 (1988); Anfossi et al., Proc. Natl. Acad. Sci. 86:3379 (1989)). However, the skilled artisan would appreciate the present invention's usefulness would not be limited to treatment of proliferative diseases, disorders, and/or conditions of
- 25 hematopoietic cells and tissues, in light of the numerous cells and cell types of varying origins which are known to exhibit proliferative phenotypes.

In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Antisense techniques are discussed, for example, in Okano, J. Neurochem. 56: 560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). Both methods rely on binding of the polynucleotide to a complementary DNA or RNA. For these techniques, preferred polynucleotides are usually oligonucleotides 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991) ) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxy-nucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat or prevent disease.

Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell.

The polynucleotides are also useful for identifying individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more  
5 restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

10 The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, individuals can be identified because each individual will have a  
15 unique set of DNA sequences. Once an unique ID database is established for an individual, positive identification of that individual, living or dead, can be made from extremely small tissue samples.

Forensic biology also benefits from using DNA-based identification techniques as disclosed herein. DNA sequences taken from very small biological  
20 samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, semen, synovial fluid, amniotic fluid, breast milk, lymph, pulmonary sputum or surfactant, urine, fecal matter, etc., can be amplified using PCR. In one prior art technique, gene sequences amplified from polymorphic loci, such as DQa class II HLA gene, are used in forensic biology to identify individuals. (Erlich, H., PCR  
25 Technology, Freeman and Co. (1992).) Once these specific polymorphic loci are amplified, they are digested with one or more restriction enzymes, yielding an



identifying set of bands on a Southern blot probed with DNA corresponding to the DQa class II HLA gene. Similarly, polynucleotides of the present invention can be used as polymorphic markers for forensic purposes.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination.

In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

### **Uses of the Polypeptides**

Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell . Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene

expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine ( $^{125}\text{I}$ ,  $^{121}\text{I}$ ), carbon ( $^{14}\text{C}$ ), sulfur ( $^{35}\text{S}$ ), tritium ( $^3\text{H}$ ), indium ( $^{112}\text{In}$ ), and  
5 technetium ( $^{99\text{m}}\text{Tc}$ ), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

In addition to assaying secreted protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-  
10 radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

15 A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example,  $^{131}\text{I}$ ,  $^{112}\text{In}$ ,  $^{99\text{m}}\text{Tc}$ ), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the  
20 subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of  $^{99\text{m}}\text{Tc}$ . The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein.  
25 In vivo tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging:

The Radiochemical Detection of Cancer, S.W. Burchiel and B. A. Rhodes, eds.,  
Masson Publishing Inc. (1982).)

Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells  
5 or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a  
10 predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

15 Moreover, polypeptides of the present invention can be used to treat, prevent, and/or diagnose disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B, SOD, catalase, DNA repair  
20 proteins), to inhibit the activity of a polypeptide (e.g., an oncogene or tumor suppressor), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth inhibition, enhancement of the immune  
25 response to proliferative cells or tissues).

Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat, prevent, and/or diagnose disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can  
5 activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also  
10 be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

### 15 **Gene Therapy Methods**

Another aspect of the present invention is to gene therapy methods for treating or preventing disorders, diseases and conditions. The gene therapy methods relate to the introduction of nucleic acid (DNA, RNA and antisense DNA or RNA) sequences into an animal to achieve expression of a polypeptide of the present invention. This  
20 method requires a polynucleotide which codes for a polypeptide of the invention that operatively linked to a promoter and any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques are known in the art, see, for example, WO90/11092, which is herein incorporated by reference.

25 Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) comprising a promoter operably linked to a

polynucleotide of the invention *ex vivo*, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, see Belldegrün et al., J. Natl. Cancer Inst., 85:207-216 (1993); Ferrantini et al., Cancer Research, 53:107-1112 (1993); Ferrantini et al., J. Immunology 153: 4604-4615 (1994); Kaido, T., et al., Int. J. Cancer 60: 221-229 (1995); Ogura et al., Cancer Research 50: 5102-5106 (1990); Santodonato, et al., Human Gene Therapy 7:1-10 (1996); Santodonato, et al., Gene Therapy 4:1246-1255 (1997); and Zhang, et al., Cancer Gene Therapy 3: 31-38 (1996)), which are herein incorporated by reference. In one embodiment, the cells which are engineered are arterial cells. The arterial cells may be reintroduced into the patient through direct injection to the artery, the tissues surrounding the artery, or through catheter injection.

As discussed in more detail below, the polynucleotide constructs can be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, and the like). The polynucleotide constructs may be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

In one embodiment, the polynucleotide of the invention is delivered as a naked polynucleotide. The term "naked" polynucleotide, DNA or RNA refers to sequences that are free from any delivery vehicle that acts to assist, promote or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the invention can also be delivered in liposome formulations and lipofectin formulations and the like can be prepared by methods well known to those skilled in the art. Such methods are described, for example, in U.S. Patent Nos. 5,593,972, 5,589,466, and 5,580,859, which are herein incorporated by reference.

The polynucleotide vector constructs of the invention used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Appropriate vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; 5 pSVK3, pBPV, pMSG and pSVL available from Pharmacia; and pEF1/V5, pcDNA3.1, and pRc/CMV2 available from Invitrogen. Other suitable vectors will be readily apparent to the skilled artisan.

Any strong promoter known to those skilled in the art can be used for driving the expression of polynucleotide sequence of the invention. Suitable promoters 10 include adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoA1 promoter; human globin promoters; viral thymidine kinase promoters, 15 such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the b-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter for the polynucleotides of the invention.

Unlike other gene therapy techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the 20 polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct of the invention can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, 25 spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye,

gland, and connective tissue. Interstitial space of the tissues comprises the intercellular, fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked *nucleic acid* sequence injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 mg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration.

The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked DNA constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The naked polynucleotides are delivered by any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, and so-called "gene guns". These delivery methods are known in the art.

5       The constructs may also be delivered with delivery vehicles such as viral sequences, viral particles, liposome formulations, lipofectin, precipitating agents, etc. Such methods of delivery are known in the art.

In certain embodiments, the polynucleotide constructs of the invention are complexed in a liposome preparation. Liposomal preparations for use in the instant  
10       invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner et al., Proc. Natl. Acad. Sci. USA ,  
15       84:7413-7416 (1987), which is herein incorporated by reference); mRNA (Malone et al., Proc. Natl. Acad. Sci. USA , 86:6077-6081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs et al., J. Biol. Chem., 265:10189-10192 (1990), which is herein incorporated by reference), in functional form.

20       Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner et al., Proc. Natl Acad. Sci. USA , 84:7413-7416 (1987), which is herein incorporated by reference). Other commercially  
25       available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).



Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication NO: WO 90/11092 (which is herein incorporated by reference) for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation  
5 of DOTMA liposomes is explained in the literature, see, e.g., Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7417, which is herein incorporated by reference. Similar methods can be used to prepare liposomes from other cationic lipid materials.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily  
10 available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these  
15 materials are well known in the art.

For example, commercially dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be  
20 prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at  
25 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to

produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred.

- 5 The various liposome-nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger et al., *Methods of Immunology* , 101:512-527 (1983), which is herein incorporated by reference. For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated.
- 10 SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and
- 15 then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include  $\text{Ca}^{2+}$ -EDTA chelation (Papahadjopoulos et al., *Biochim. Biophys. Acta*, 394:483 (1975); Wilson et al., *Cell* , 17:77 (1979)); ether injection (Deamer et al., *Biochim. Biophys. Acta*, 443:629 (1976); Ostro et al., *Biochem. Biophys. Res. Commun.*, 76:836 (1977); Fraley et al., *Proc. Natl. Acad. Sci. USA*, 76:3348 (1979)); detergent dialysis (Enoch et al., *Proc. Natl. Acad. Sci. USA* , 76:145 (1979)); and reverse-phase evaporation (REV) (Fraley et al., *J. Biol. Chem.*, 255:10431 (1980);
- 25 Szoka et al., *Proc. Natl. Acad. Sci. USA* , 75:145 (1978); Schaefer-Ridder et al., *Science*, 215:166 (1982)), which are herein incorporated by reference.

Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

U.S. Patent NO: 5,676,954 (which is herein incorporated by reference) reports  
5 on the injection of genetic material, complexed with cationic liposomes carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication NO: WO 94/9469 (which are herein incorporated by reference) provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622,  
10 5,580,859, 5,703,055, and international publication NO: WO 94/9469 (which are herein incorporated by reference) provide methods for delivering DNA-cationic lipid complexes to mammals.

In certain embodiments, cells are engineered, *ex vivo* or *in vivo*, using a retroviral particle containing RNA which comprises a sequence encoding  
15 polypeptides of the invention. Retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, Rous sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

20 The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, R-2, R-AM, PA12, T19-14X, VT-19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, Human Gene Therapy , 1:5-14 (1990), which is incorporated herein by  
25 reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation,

the use of liposomes, and  $\text{CaPO}_4$  precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which  
5 include polynucleotide encoding polypeptides of the invention. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either *in vitro* or *in vivo*. The transduced eukaryotic cells will express polypeptides of the invention.

In certain other embodiments, cells are engineered, *ex vivo* or *in vivo*, with polynucleotides of the invention contained in an adenovirus vector. Adenovirus can  
10 be manipulated such that it encodes and expresses polypeptides of the invention, and at the same time is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. Adenovirus expression is achieved without integration of the viral DNA into the host cell chromosome, thereby alleviating concerns about insertional mutagenesis. Furthermore, adenoviruses have been used as live enteric vaccines for  
15 many years with an excellent safety profile (Schwartz et al., Am. Rev. Respir. Dis., 109:233-238 (1974)). Finally, adenovirus mediated gene transfer has been demonstrated in a number of instances including transfer of alpha-1-antitrypsin and CFTR to the lungs of cotton rats (Rosenfeld et al., Science, 252:431-434 (1991); Rosenfeld et al., Cell, 68:143-155 (1992)). Furthermore, extensive studies to attempt  
20 to establish adenovirus as a causative agent in human cancer were uniformly negative (Green et al. Proc. Natl. Acad. Sci. USA, 76:6606 (1979)).

Suitable adenoviral vectors useful in the present invention are described, for example, in Kozarsky and Wilson, Curr. Opin. Genet. Devel., 3:499-503 (1993); Rosenfeld et al., Cell, 68:143-155 (1992); Engelhardt et al., Human Genet. Ther.,  
25 4:759-769 (1993); Yang et al., Nature Genet., 7:362-369 (1994); Wilson et al., Nature, 365:691-692 (1993); and U.S. Patent NO: 5,652,224, which are herein

incorporated by reference. For example, the adenovirus vector Ad2 is useful and can be grown in human 293 cells. These cells contain the E1 region of adenovirus and constitutively express E1a and E1b, which complement the defective adenoviruses by providing the products of the genes deleted from the vector. In addition to Ad2, other  
5 varieties of adenovirus (e.g., Ad3, Ad5, and Ad7) are also useful in the present invention.

Preferably, the adenoviruses used in the present invention are replication deficient. Replication deficient adenoviruses require the aid of a helper virus and/or packaging cell line to form infectious particles. The resulting virus is capable of  
10 infecting cells and can express a polynucleotide of interest which is operably linked to a promoter, but cannot replicate in most cells. Replication deficient adenoviruses may be deleted in one or more of all or a portion of the following genes: E1a, E1b, E3, E4, E2a, or L1 through L5.

In certain other embodiments, the cells are engineered, *ex vivo* or *in vivo*,  
15 using an adeno-associated virus (AAV). AAVs are naturally occurring defective viruses that require helper viruses to produce infectious particles (Muzyczka, Curr. Topics in Microbiol. Immunol., 158:97 (1992)). It is also one of the few viruses that may integrate its DNA into non-dividing cells. Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate, but space for exogenous DNA is  
20 limited to about 4.5 kb. Methods for producing and using such AAVs are known in the art. See, for example, U.S. Patent Nos. 5,139,941, 5,173,414, 5,354,678, 5,436,146, 5,474,935, 5,478,745, and 5,589,377.

For example, an appropriate AAV vector for use in the present invention will include all the sequences necessary for DNA replication, encapsidation, and host-cell  
25 integration. The polynucleotide construct containing polynucleotides of the invention is inserted into the AAV vector using standard cloning methods, such as those found

in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (1989). The recombinant AAV vector is then transfected into packaging cells which are infected with a helper virus, using any standard technique, including lipofection, electroporation, calcium phosphate precipitation, etc. Appropriate helper  
5 viruses include adenoviruses, cytomegaloviruses, vaccinia viruses, or herpes viruses. Once the packaging cells are transfected and infected, they will produce infectious AAV viral particles which contain the polynucleotide construct of the invention. These viral particles are then used to transduce eukaryotic cells, either *ex vivo* or *in vivo*. The transduced cells will contain the polynucleotide construct integrated into its  
10 genome, and will express the desired gene product.

Another method of gene therapy involves operably associating heterologous control regions and endogenous polynucleotide sequences (e.g. encoding the polypeptide sequence of interest) via homologous recombination (see, e.g., U.S. Patent NO: 5,641,670, issued June 24, 1997; International Publication NO: WO  
15 96/29411, published September 26, 1996; International Publication NO: WO 94/12650, published August 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA, 86:8932-8935 (1989); and Zijlstra et al., Nature, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not normally expressed in the cells, or is expressed at a lower level than desired.

20 Polynucleotide constructs are made, using standard techniques known in the art, which contain the promoter with targeting sequences flanking the promoter. Suitable promoters are described herein. The targeting sequence is sufficiently complementary to an endogenous sequence to permit homologous recombination of the promoter-targeting sequence with the endogenous sequence. The targeting  
25 sequence will be sufficiently near the 5' end of the desired endogenous

polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination.

The promoter and the targeting sequences can be amplified using PCR.

Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter. The amplified promoter and targeting sequences are digested and ligated together.

The promoter-targeting sequence construct is delivered to the cells, either as naked polynucleotide, or in conjunction with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, whole viruses, lipofection, precipitating agents, etc., described in more detail above. The P promoter-targeting sequence can be delivered by any method, included direct needle injection, intravenous injection, topical administration, catheter infusion, particle accelerators, etc. The methods are described in more detail below.

The promoter-targeting sequence construct is taken up by cells. Homologous recombination between the construct and the endogenous sequence takes place, such that an endogenous sequence is placed under the control of the promoter. The promoter then drives the expression of the endogenous sequence.

The polynucleotides encoding polypeptides of the present invention may be administered along with other polynucleotides encoding other angiogenic proteins. Angiogenic proteins include, but are not limited to, acidic and basic fibroblast growth factors, VEGF-1, VEGF-2 (VEGF-C), VEGF-3 (VEGF-B), epidermal growth factor alpha and beta, platelet-derived endothelial cell growth factor, platelet-derived growth factor, tumor necrosis factor alpha, hepatocyte growth factor, insulin like growth

factor, colony stimulating factor, macrophage colony stimulating factor, granulocyte/macrophage colony stimulating factor, and nitric oxide synthase.

Preferably, the polynucleotide encoding a polypeptide of the invention contains a secretory signal sequence that facilitates secretion of the protein.

5 Typically, the signal sequence is positioned in the coding region of the polynucleotide to be expressed towards or at the 5' end of the coding region. The signal sequence may be homologous or heterologous to the polynucleotide of interest and may be homologous or heterologous to the cells to be transfected. Additionally, the signal sequence may be chemically synthesized using methods known in the art.

10 Any mode of administration of any of the above-described polynucleotides constructs can be used so long as the mode results in the expression of one or more molecules in an amount sufficient to provide a therapeutic effect. This includes direct needle injection, systemic injection, catheter infusion, biolistic injectors, particle accelerators (i.e., "gene guns"), gelfoam sponge depots, other commercially available  
15 depot materials, osmotic pumps (e.g., Alza minipumps), oral or suppositorial solid (tablet or pill) pharmaceutical formulations, and decanting or topical applications during surgery. For example, direct injection of naked calcium phosphate-precipitated plasmid into rat liver and rat spleen or a protein-coated plasmid into the portal vein has resulted in gene expression of the foreign gene in the  
20 rat livers. (Kaneda et al., Science, 243:375 (1989)).

A preferred method of local administration is by direct injection. Preferably, a recombinant molecule of the present invention complexed with a delivery vehicle is administered by direct injection into or locally within the area of arteries.

Administration of a composition locally within the area of arteries refers to injecting  
25 the composition centimeters and preferably, millimeters within arteries.



Another method of local administration is to contact a polynucleotide construct of the present invention in or around a surgical wound. For example, a patient can undergo surgery and the polynucleotide construct can be coated on the surface of tissue inside the wound or the construct can be injected into areas of tissue  
5 inside the wound.

Therapeutic compositions useful in systemic administration, include recombinant molecules of the present invention complexed to a targeted delivery vehicle of the present invention. Suitable delivery vehicles for use with systemic administration comprise liposomes comprising ligands for targeting the vehicle to a  
10 particular site.

Preferred methods of systemic administration, include intravenous injection, aerosol, oral and percutaneous (topical) delivery. Intravenous injections can be performed using methods standard in the art. Aerosol delivery can also be performed using methods standard in the art (see, for example, Stribling et al., Proc. Natl. Acad. Sci. USA , 189:11277-11281 (1992), which is incorporated herein by reference). Oral  
15 delivery can be performed by complexing a polynucleotide construct of the present invention to a carrier capable of withstanding degradation by digestive enzymes in the gut of an animal. Examples of such carriers, include plastic capsules or tablets, such as those known in the art. Topical delivery can be performed by mixing a  
20 polynucleotide construct of the present invention with a lipophilic reagent (e.g., DMSO) that is capable of passing into the skin.

Determining an effective amount of substance to be delivered can depend upon a number of factors including, for example, the chemical structure and biological activity of the substance, the age and weight of the animal, the precise  
25 condition requiring treatment and its severity, and the route of administration. The frequency of treatments depends upon a number of factors, such as the amount of

polynucleotide constructs administered per dose, as well as the health and history of the subject. The precise amount, number of doses, and timing of doses will be determined by the attending physician or veterinarian. Therapeutic compositions of the present invention can be administered to any animal, preferably to mammals and  
5 birds. Preferred mammals include humans, dogs, cats, mice, rats, rabbits sheep, cattle, horses and pigs, with humans being particularly

### **Biological Activities**

The polynucleotides or polypeptides, or agonists or antagonists of the present  
10 invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the diseases associated with the biological activity. Thus, the polynucleotides or polypeptides, or agonists or antagonists could be used to treat the associated disease.

15

### **Immune Activity**

The polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders, and/or conditions of the immune system, by activating or inhibiting the proliferation,  
20 differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune diseases, disorders, and/or conditions may be genetic, somatic, such as cancer or some autoimmune diseases,  
25 disorders, and/or conditions, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present

invention can be used as a marker or detector of a particular immune system disease or disorder.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders, and/or conditions of hematopoietic cells. A polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat or prevent those diseases, disorders, and/or conditions associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein diseases, disorders, and/or conditions (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to treat or prevent blood coagulation diseases, disorders, and/or conditions (e.g., afibrinogenemia, factor deficiencies), blood platelet diseases, disorders, and/or conditions (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention that can decrease hemostatic or thrombolytic activity could be used

to inhibit or dissolve clotting. These molecules could be important in the treatment or prevention of heart attacks (infarction), strokes, or scarring.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be useful in treating, preventing, and/or diagnosing autoimmune diseases, disorders, and/or conditions. Many autoimmune diseases, disorders, and/or conditions result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune diseases, disorders, and/or conditions.

Examples of autoimmune diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or agonists or antagonists of the

present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to treat, prevent, and/or diagnose organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

Similarly, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide or agonists or antagonist may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat, prevent, and/or diagnose inflammatory conditions, both chronic and acute conditions, including chronic prostatitis, granulomatous prostatitis and malacoplakia, inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

## **Hyperproliferative Disorders**

A polynucleotides or polypeptides, or agonists or antagonists of the invention can be used to treat, prevent, and/or diagnose hyperproliferative diseases, disorders, including neoplasms. A polynucleotides or polypeptides, or agonists or antagonists of the present invention may inhibit the proliferation of the disorder through direct or  
5 indirect interactions. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating,  
10 differentiating, or mobilizing T-cells, hyperproliferative diseases, disorders, and/or conditions can be treated, prevented, and/or diagnosed. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating, preventing, and/or diagnosing hyperproliferative diseases,  
15 disorders, and/or conditions, such as a chemotherapeutic agent.

Examples of hyperproliferative diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or agonists or antagonists of the present invention include, but are not limited to neoplasms located in the: colon, abdomen, bone, breast, digestive system, liver,  
20 pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

Similarly, other hyperproliferative diseases, disorders, and/or conditions can also be treated, prevented, and/or diagnosed by a polynucleotides or polypeptides, or  
25 agonists or antagonists of the present invention. Examples of such hyperproliferative diseases, disorders, and/or conditions include, but are not limited to:

hypergammaglobulinemia, lymphoproliferative diseases, disorders, and/or conditions, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstrom's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

5           One preferred embodiment utilizes polynucleotides of the present invention to inhibit aberrant cellular division, by gene therapy using the present invention, and/or protein fusions or fragments thereof.

          Thus, the present invention provides a method for treating or preventing cell proliferative diseases, disorders, and/or conditions by inserting into an abnormally  
10   proliferating cell a polynucleotide of the present invention, wherein said polynucleotide represses said expression.

          Another embodiment of the present invention provides a method of treating or preventing cell-proliferative diseases, disorders, and/or conditions in individuals comprising administration of one or more active gene copies of the present invention:  
15   to an abnormally proliferating cell or cells. In a preferred embodiment, polynucleotides of the present invention is a DNA construct comprising a recombinant expression vector effective in expressing a DNA sequence encoding said polynucleotides. In another preferred embodiment of the present invention, the DNA construct encoding the polynucleotides of the present invention is inserted into cells to  
20   be treated utilizing a retrovirus, or more preferably an adenoviral vector (See G J. Nabel, et. al., PNAS 1999 96: 324-326, which is hereby incorporated by reference). In a most preferred embodiment, the viral vector is defective and will not transform non-proliferating cells, only proliferating cells. Moreover, in a preferred  
25   embodiment, the polynucleotides of the present invention inserted into proliferating cells either alone, or in combination with or fused to other polynucleotides, can then be modulated via an external stimulus (i.e. magnetic, specific small molecule,

chemical, or drug administration, etc.), which acts upon the promoter upstream of said polynucleotides to induce expression of the encoded protein product. As such the beneficial therapeutic affect of the present invention may be expressly modulated (i.e. to increase, decrease, or inhibit expression of the present invention) based upon said  
5 external stimulus.

Polynucleotides of the present invention may be useful in repressing expression of oncogenic genes or antigens. By "repressing expression of the oncogenic genes " is intended the suppression of the transcription of the gene, the degradation of the gene transcript (pre-message RNA), the inhibition of splicing, the  
10 destruction of the messenger RNA, the prevention of the post-translational modifications of the protein, the destruction of the protein, or the inhibition of the normal function of the protein.

For local administration to abnormally proliferating cells, polynucleotides of the present invention may be administered by any method known to those of skill in  
15 the art including, but not limited to transfection, electroporation, microinjection of cells, or in vehicles such as liposomes, lipofectin, or as naked polynucleotides, or any other method described throughout the specification. The polynucleotide of the present invention may be delivered by known gene delivery systems such as, but not limited to, retroviral vectors (Gilboa, J. Virology 44:845 (1982); Hocke, Nature  
20 320:275 (1986); Wilson, et al., Proc. Natl. Acad. Sci. U.S.A. 85:3014), vaccinia virus system (Chakrabarty et al., Mol. Cell Biol. 5:3403 (1985) or other efficient DNA delivery systems (Yates et al., Nature 313:812 (1985)) known to those skilled in the art. These references are exemplary only and are hereby incorporated by reference. In order to specifically deliver or transfect cells which are abnormally proliferating  
25 and spare non-dividing cells, it is preferable to utilize a retrovirus, or adenoviral (as described in the art and elsewhere herein) delivery system known to those of skill in



the art. Since host DNA replication is required for retroviral DNA to integrate and the retrovirus will be unable to self replicate due to the lack of the retrovirus genes needed for its life cycle. Utilizing such a retroviral delivery system for polynucleotides of the present invention will target said gene and constructs to

5 abnormally proliferating cells and will spare the non-dividing normal cells.

The polynucleotides of the present invention may be delivered directly to cell proliferative disorder/disease sites in internal organs, body cavities and the like by use of imaging devices used to guide an injecting needle directly to the disease site. The polynucleotides of the present invention may also be administered to disease sites at

10 the time of surgical intervention.

By "cell proliferative disease" is meant any human or animal disease or disorder, affecting any one or any combination of organs, cavities, or body parts, which is characterized by single or multiple local abnormal proliferations of cells, groups of cells, or tissues, whether benign or malignant.

15 Any amount of the polynucleotides of the present invention may be administered as long as it has a biologically inhibiting effect on the proliferation of the treated cells. Moreover, it is possible to administer more than one of the polynucleotide of the present invention simultaneously to the same site. By "biologically inhibiting" is meant partial or total growth inhibition as well as

20 decreases in the rate of proliferation or growth of the cells. The biologically inhibitory dose may be determined by assessing the effects of the polynucleotides of the present invention on target malignant or abnormally proliferating cell growth in tissue culture, tumor growth in animals and cell cultures, or any other method known to one of ordinary skill in the art.

25 The present invention is further directed to antibody-based therapies which involve administering of anti-polypeptides and anti-polynucleotide antibodies to a

mammalian, preferably human, patient for treating, preventing, and/or diagnosing one or more of the described diseases, disorders, and/or conditions. Methods for producing anti-polypeptides and anti-polynucleotide antibodies polyclonal and monoclonal antibodies are described in detail elsewhere herein. Such antibodies may  
5 be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the  
10 antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

15 In particular, the antibodies, fragments and derivatives of the present invention are useful for treating, preventing, and/or diagnosing a subject having or developing cell proliferative and/or differentiation diseases, disorders, and/or conditions as described herein. Such treatment comprises administering a single or multiple doses of the antibody, or a fragment, derivative, or a conjugate thereof.

20 The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors, for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

It is preferred to use high affinity and/or potent in vivo inhibiting and/or  
25 neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and

therapy of diseases, disorders, and/or conditions related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides, including fragments thereof. Preferred binding affinities include those  
5 with a dissociation constant or  $K_d$  less than  $5 \times 10^{-6}M$ ,  $10^{-6}M$ ,  $5 \times 10^{-7}M$ ,  $10^{-7}M$ ,  $5 \times 10^{-8}M$ ,  $10^{-8}M$ ,  $5 \times 10^{-9}M$ ,  $10^{-9}M$ ,  $5 \times 10^{-10}M$ ,  $10^{-10}M$ ,  $5 \times 10^{-11}M$ ,  $10^{-11}M$ ,  $5 \times 10^{-12}M$ ,  $10^{-12}M$ ,  $5 \times 10^{-13}M$ ,  $10^{-13}M$ ,  $5 \times 10^{-14}M$ ,  $10^{-14}M$ ,  $5 \times 10^{-15}M$ , and  $10^{-15}M$ .

Moreover, polypeptides of the present invention are useful in inhibiting the angiogenesis of proliferative cells or tissues, either alone, as a protein fusion, or in  
10 combination with other polypeptides directly or indirectly, as described elsewhere herein. In a most preferred embodiment, said anti-angiogenesis effect may be achieved indirectly, for example, through the inhibition of hematopoietic, tumor-specific cells, such as tumor-associated macrophages (See Joseph IB, et al. J Natl Cancer Inst, 90(21):1648-53 (1998), which is hereby incorporated by reference).  
15 Antibodies directed to polypeptides or polynucleotides of the present invention may also result in inhibition of angiogenesis directly, or indirectly (See Witte L, et al., Cancer Metastasis Rev. 17(2):155-61 (1998), which is hereby incorporated by reference)).

Polypeptides, including protein fusions, of the present invention, or fragments  
20 thereof may be useful in inhibiting proliferative cells or tissues through the induction of apoptosis. Said polypeptides may act either directly, or indirectly to induce apoptosis of proliferative cells and tissues, for example in the activation of a death-domain receptor, such as tumor necrosis factor (TNF) receptor-1, CD95 (Fas/APO-1), TNF-receptor-related apoptosis-mediated protein (TRAMP) and TNF-related  
25 apoptosis-inducing ligand (TRAIL) receptor-1 and -2 (See Schulze-Osthoff K, et.al., Eur J Biochem 254(3):439-59 (1998), which is hereby incorporated by reference).

Moreover, in another preferred embodiment of the present invention, said polypeptides may induce apoptosis through other mechanisms, such as in the activation of other proteins which will activate apoptosis, or through stimulating the expression of said proteins, either alone or in combination with small molecule drugs or adjuvants, such as apoptonin, galectins, thioredoxins, antiinflammatory proteins (See for example, Mutat Res 400(1-2):447-55 (1998), Med Hypotheses.50(5):423-33 (1998), Chem Biol Interact. Apr 24;111-112:23-34 (1998), J Mol Med.76(6):402-12 (1998), Int J Tissue React;20(1):3-15 (1998), which are all hereby incorporated by reference).

10 Polypeptides, including protein fusions to, or fragments thereof, of the present invention are useful in inhibiting the metastasis of proliferative cells or tissues. Inhibition may occur as a direct result of administering polypeptides, or antibodies directed to said polypeptides as described elsewhere herein, or indirectly, such as activating the expression of proteins known to inhibit metastasis, for example alpha 4:  
15 integrins, (See, e.g., Curr Top Microbiol Immunol 1998;231:125-41, which is hereby incorporated by reference). Such thereapeutic affects of the present invention may be achieved either alone, or in combination with small molecule drugs or adjuvants.

In another embodiment, the invention provides a method of delivering compositions containing the polypeptides of the invention (e.g., compositions  
20 containing polypeptides or polypeptide antibodies associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs) to targeted cells expressing the polypeptide of the present invention. Polypeptides or polypeptide antibodies of the invention may be associated with with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic  
25 and/or covalent interactions.

Polypeptides, protein fusions to, or fragments thereof, of the present invention are useful in enhancing the immunogenicity and/or antigenicity of proliferating cells or tissues, either directly, such as would occur if the polypeptides of the present invention 'vaccinated' the immune response to respond to proliferative antigens and  
5 immunogens, or indirectly, such as in activating the expression of proteins known to enhance the immune response (e.g. chemokines), to said antigens and immunogens.

### **Cardiovascular Disorders**

Polynucleotides or polypeptides, or agonists or antagonists of the invention  
10 may be used to treat, prevent, and/or diagnose cardiovascular diseases, disorders, and/or conditions, including peripheral artery disease, such as limb ischemia.

Cardiovascular diseases, disorders, and/or conditions include cardiovascular abnormalities, such as arterio-arterial fistula, arteriovenous fistula, cerebral arteriovenous malformations, congenital heart defects, pulmonary atresia, and  
15 Scimitar Syndrome. Congenital heart defects include aortic coarctation, cor triatriatum, coronary vessel anomalies, crisscross heart, dextrocardia, patent ductus arteriosus, Ebstein's anomaly, Eisenmenger complex, hypoplastic left heart syndrome, levocardia, tetralogy of fallot, transposition of great vessels, double outlet right ventricle, tricuspid atresia, persistent truncus arteriosus, and heart septal defects, such  
20 as aortopulmonary septal defect, endocardial cushion defects, Lutembacher's Syndrome, trilogly of Fallot, ventricular heart septal defects.

Cardiovascular diseases, disorders, and/or conditions also include heart disease, such as arrhythmias, carcinoid heart disease, high cardiac output, low cardiac output, cardiac tamponade, endocarditis (including bacterial), heart aneurysm, cardiac  
25 arrest, congestive heart failure, congestive cardiomyopathy, paroxysmal dyspnea, cardiac edema, heart hypertrophy, congestive cardiomyopathy, left ventricular

hypertrophy, right ventricular hypertrophy, post-infarction heart rupture, ventricular septal rupture, heart valve diseases, myocardial diseases, myocardial ischemia, pericardial effusion, pericarditis (including constrictive and tuberculous), pneumopericardium, postpericardiotomy syndrome, pulmonary heart disease, 5 rheumatic heart disease, ventricular dysfunction, hyperemia, cardiovascular pregnancy complications, Scimitar Syndrome, cardiovascular syphilis, and cardiovascular tuberculosis.

Arrhythmias include sinus arrhythmia, atrial fibrillation, atrial flutter, bradycardia, extrasystole, Adams-Stokes Syndrome, bundle-branch block, sinoatrial 10 block, long QT syndrome, parasystole, Lown-Ganong-Levine Syndrome, Mahaim-type pre-excitation syndrome, Wolff-Parkinson-White syndrome, sick sinus syndrome, tachycardias, and ventricular fibrillation. Tachycardias include paroxysmal tachycardia, supraventricular tachycardia, accelerated idioventricular rhythm, atrioventricular nodal reentry tachycardia, ectopic atrial tachycardia, ectopic 15 junctional tachycardia, sinoatrial nodal reentry tachycardia, sinus tachycardia, Torsades de Pointes, and ventricular tachycardia.

Heart valve disease include aortic valve insufficiency, aortic valve stenosis, hear murmurs, aortic valve prolapse, mitral valve prolapse, tricuspid valve prolapse, mitral valve insufficiency, mitral valve stenosis, pulmonary atresia, pulmonary valve 20 insufficiency, pulmonary valve stenosis, tricuspid atresia, tricuspid valve insufficiency, and tricuspid valve stenosis.

Myocardial diseases include alcoholic cardiomyopathy, congestive cardiomyopathy, hypertrophic cardiomyopathy, aortic subvalvular stenosis, pulmonary subvalvular stenosis, restrictive cardiomyopathy, Chagas cardiomyopathy, 25 endocardial fibroelastosis, endomyocardial fibrosis, Kearns Syndrome, myocardial reperfusion injury, and myocarditis.

Myocardial ischemias include coronary disease, such as angina pectoris, coronary aneurysm, coronary arteriosclerosis, coronary thrombosis, coronary vasospasm, myocardial infarction and myocardial stunning.

Cardiovascular diseases also include vascular diseases such as aneurysms, angiodyplasia, angiomas, bacillary angiomas, Hippel-Lindau Disease, Klippel-Trenaunay-Weber Syndrome, Sturge-Weber Syndrome, angioneurotic edema, aortic diseases, Takayasu's Arteritis, aortitis, Leriche's Syndrome, arterial occlusive diseases, arteritis, enarteritis, polyarteritis nodosa, cerebrovascular diseases, disorders, and/or conditions, diabetic angiopathies, diabetic retinopathy, embolisms, thrombosis, erythromelalgia, hemorrhoids, hepatic veno-occlusive disease, hypertension, hypotension, ischemia, peripheral vascular diseases, phlebitis, pulmonary veno-occlusive disease, Raynaud's disease, CREST syndrome, retinal vein occlusion, Scimitar syndrome, superior vena cava syndrome, telangiectasia, ataxia telangiectasia, hereditary hemorrhagic telangiectasia, varicocele, varicose veins, varicose ulcer, vasculitis, and venous insufficiency.

Aneurysms include dissecting aneurysms, false aneurysms, infected aneurysms, ruptured aneurysms, aortic aneurysms, cerebral aneurysms, coronary aneurysms, heart aneurysms, and iliac aneurysms.

Arterial occlusive diseases include arteriosclerosis, intermittent claudication, carotid stenosis, fibromuscular dysplasias, mesenteric vascular occlusion, Moyamoya disease, renal artery obstruction, retinal artery occlusion, and thromboangiitis obliterans.

Cerebrovascular diseases, disorders, and/or conditions include carotid artery diseases, cerebral amyloid angiopathy, cerebral aneurysm, cerebral anoxia, cerebral arteriosclerosis, cerebral arteriovenous malformation, cerebral artery diseases, cerebral embolism and thrombosis, carotid artery thrombosis, sinus thrombosis,

Wallenberg's syndrome, cerebral hemorrhage, epidural hematoma, subdural hematoma, subarachnoid hemorrhage, cerebral infarction, cerebral ischemia (including transient), subclavian steal syndrome, periventricular leukomalacia, vascular headache, cluster headache, migraine, and vertebrobasilar insufficiency.

5            Embolisms include air embolisms, amniotic fluid embolisms, cholesterol embolisms, blue toe syndrome, fat embolisms, pulmonary embolisms, and thromboembolisms. Thrombosis include coronary thrombosis, hepatic vein thrombosis, retinal vein occlusion, carotid artery thrombosis, sinus thrombosis, Wallenberg's syndrome, and thrombophlebitis.

10           Ischemia includes cerebral ischemia, ischemic colitis, compartment syndromes, anterior compartment syndrome, myocardial ischemia, reperfusion injuries, and peripheral limb ischemia. Vasculitis includes aortitis, arteritis, Behcet's Syndrome, Churg-Strauss Syndrome, mucocutaneous lymph node syndrome, thromboangiitis obliterans, hypersensitivity vasculitis, Schoenlein-Henoch purpura,  
15           allergic cutaneous vasculitis, and Wegener's granulomatosis.

             Polynucleotides or polypeptides, or agonists or antagonists of the invention, are especially effective for the treatment of critical limb ischemia and coronary disease.

             Polypeptides may be administered using any method known in the art,  
20           including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, biolistic injectors, particle accelerators, gelfoam sponge depots, other commercially available depot materials, osmotic pumps, oral or suppositorial solid pharmaceutical formulations, decanting or topical applications during surgery, aerosol delivery. Such methods are known in the  
25           art. Polypeptides of the invention may be administered as part of a *Therapeutic*,



described in more detail below. Methods of delivering polynucleotides of the invention are described in more detail herein.

### **Anti-Angiogenesis Activity**

5           The naturally occurring balance between endogenous stimulators and inhibitors of angiogenesis is one in which inhibitory influences predominate. Rastinejad *et al.*, *Cell* 56:345-355 (1989). In those rare instances in which neovascularization occurs under normal physiological conditions, such as wound healing, organ regeneration, embryonic development, and female reproductive  
10       processes, angiogenesis is stringently regulated and spatially and temporally delimited. Under conditions of pathological angiogenesis such as that characterizing solid tumor growth, these regulatory controls fail. Unregulated angiogenesis becomes pathologic and sustains progression of many neoplastic and non-neoplastic diseases. A number of serious diseases are dominated by abnormal neovascularization  
15       including solid tumor growth and metastases, arthritis, some types of eye diseases, disorders, and/or conditions, and psoriasis. See, e.g., reviews by Moses *et al.*, *Biotech.* 9:630-634 (1991); Folkman *et al.*, *N. Engl. J. Med.*, 333:1757-1763 (1995); Auerbach *et al.*, *J. Microvasc. Res.* 29:401-411 (1985); Folkman, *Advances in Cancer Research*, eds. Klein and Weinhouse, Academic Press, New York, pp. 175-  
20       203 (1985); Patz, *Am. J. Ophthalmol.* 94:715-743 (1982); and Folkman *et al.*, *Science* 221:719-725 (1983). In a number of pathological conditions, the process of angiogenesis contributes to the disease state. For example, significant data have accumulated which suggest that the growth of solid tumors is dependent on angiogenesis. Folkman and Klagsbrun, *Science* 235:442-447 (1987).

25           The present invention provides for treatment of diseases, disorders, and/or conditions associated with neovascularization by administration of the

polynucleotides and/or polypeptides of the invention, as well as agonists or antagonists of the present invention. Malignant and metastatic conditions which can be treated with the polynucleotides and polypeptides, or agonists or antagonists of the invention include, but are not limited to, malignancies, solid tumors, and cancers  
5 described herein and otherwise known in the art (for a review of such disorders, see Fishman *et al.*, Medicine, 2d Ed., J. B. Lippincott Co., Philadelphia (1985)). Thus, the present invention provides a method of treating, preventing, and/or diagnosing an angiogenesis-related disease and/or disorder, comprising administering to an individual in need thereof a therapeutically effective amount of a polynucleotide,  
10 polypeptide, antagonist and/or agonist of the invention. For example, polynucleotides, polypeptides, antagonists and/or agonists may be utilized in a variety of additional methods in order to therapeutically treat or prevent a cancer or tumor. Cancers which may be treated, prevented, and/or diagnosed with polynucleotides, polypeptides, antagonists and/or agonists include, but are not limited to solid tumors,  
15 including prostate, lung, breast, ovarian, stomach, pancreas, larynx, esophagus, testes, liver, parotid, biliary tract, colon, rectum, cervix, uterus, endometrium, kidney, bladder, thyroid cancer; primary tumors and metastases; melanomas; glioblastoma; Kaposi's sarcoma; leiomyosarcoma; non- small cell lung cancer; colorectal cancer; advanced malignancies; and blood born tumors such as leukemias. For example,  
20 polynucleotides, polypeptides, antagonists and/or agonists may be delivered topically, in order to treat or prevent cancers such as skin cancer, head and neck tumors, breast tumors, and Kaposi's sarcoma.

Within yet other aspects, polynucleotides, polypeptides, antagonists and/or agonists may be utilized to treat superficial forms of bladder cancer by, for example,  
25 intravesical administration. Polynucleotides, polypeptides, antagonists and/or agonists may be delivered directly into the tumor, or near the tumor site, via injection or a

catheter. Of course, as the artisan of ordinary skill will appreciate, the appropriate mode of administration will vary according to the cancer to be treated. Other modes of delivery are discussed herein.

Polynucleotides, polypeptides, antagonists and/or agonists may be useful in  
5 treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions, besides cancers, which involve angiogenesis. These diseases, disorders, and/or conditions include, but are not limited to: benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas; artherosclerotic plaques; ocular angiogenic diseases, for example, diabetic retinopathy,  
10 retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, uveitis and Pterygia (abnormal blood vessel growth) of the eye; rheumatoid arthritis; psoriasis; delayed wound healing; endometriosis; vasculogenesis; granulations; hypertrophic scars (keloids); nonunion fractures; scleroderma; trachoma; vascular adhesions; myocardial,  
15 angiogenesis; coronary collaterals; cerebral collaterals; arteriovenous malformations; ischemic limb angiogenesis; Osler-Webber Syndrome; plaque neovascularization; telangiectasia; hemophilic joints; angiofibroma; fibromuscular dysplasia; wound granulation; Crohn's disease; and atherosclerosis.

For example, within one aspect of the present invention methods are provided  
20 for treating, preventing, and/or diagnosing hypertrophic scars and keloids, comprising the step of administering a polynucleotide, polypeptide, antagonist and/or agonist of the invention to a hypertrophic scar or keloid.

Within one embodiment of the present invention polynucleotides, polypeptides, antagonists and/or agonists are directly injected into a hypertrophic scar  
25 or keloid, in order to prevent the progression of these lesions. This therapy is of particular value in the prophylactic treatment of conditions which are known to result

in the development of hypertrophic scars and keloids (e.g., burns), and is preferably initiated after the proliferative phase has had time to progress (approximately 14 days after the initial injury), but before hypertrophic scar or keloid development. As noted above, the present invention also provides methods for treating, preventing, and/or  
5 diagnosing neovascular diseases of the eye, including for example, corneal neovascularization, neovascular glaucoma, proliferative diabetic retinopathy, retrolental fibroplasia and macular degeneration.

Moreover, Ocular diseases, disorders, and/or conditions associated with neovascularization which can be treated, prevented, and/or diagnosed with the  
10 polynucleotides and polypeptides of the present invention (including agonists and/or antagonists) include, but are not limited to: neovascular glaucoma, diabetic retinopathy, retinoblastoma, retrolental fibroplasia, uveitis, retinopathy of prematurity macular degeneration, corneal graft neovascularization, as well as other eye inflammatory diseases, ocular tumors and diseases associated with choroidal or iris  
15 neovascularization. See, e.g., reviews by Waltman *et al.*, *Am. J. Ophthalm.* 85:704-710 (1978) and Gartner *et al.*, *Surv. Ophthalm.* 22:291-312 (1978).

Thus, within one aspect of the present invention methods are provided for treating or preventing neovascular diseases of the eye such as corneal neovascularization (including corneal graft neovascularization), comprising the step  
20 of administering to a patient a therapeutically effective amount of a compound (as described above) to the cornea, such that the formation of blood vessels is inhibited. Briefly, the cornea is a tissue which normally lacks blood vessels. In certain pathological conditions however, capillaries may extend into the cornea from the pericorneal vascular plexus of the limbus. When the cornea becomes vascularized, it  
25 also becomes clouded, resulting in a decline in the patient's visual acuity. Visual loss may become complete if the cornea completely opacitates. A wide variety of

diseases, disorders, and/or conditions can result in corneal neovascularization, including for example, corneal infections (e.g., trachoma, herpes simplex keratitis, leishmaniasis and onchocerciasis), immunological processes (e.g., graft rejection and Stevens-Johnson's syndrome), alkali burns, trauma, inflammation (of any cause),  
5 toxic and nutritional deficiency states, and as a complication of wearing contact lenses.

Within particularly preferred embodiments of the invention, may be prepared for topical administration in saline (combined with any of the preservatives and antimicrobial agents commonly used in ocular preparations), and administered in  
10 eyedrop form. The solution or suspension may be prepared in its pure form and administered several times daily. Alternatively, anti-angiogenic compositions, prepared as described above, may also be administered directly to the cornea. Within preferred embodiments, the anti-angiogenic composition is prepared with a muco-adhesive polymer which binds to cornea. Within further embodiments, the anti-  
15 angiogenic factors or anti-angiogenic compositions may be utilized as an adjunct to conventional steroid therapy. Topical therapy may also be useful prophylactically in corneal lesions which are known to have a high probability of inducing an angiogenic response (such as chemical burns). In these instances the treatment, likely in combination with steroids, may be instituted immediately to help prevent subsequent  
20 complications.

Within other embodiments, the compounds described above may be injected directly into the corneal stroma by an ophthalmologist under microscopic guidance. The preferred site of injection may vary with the morphology of the individual lesion, but the goal of the administration would be to place the composition at the advancing  
25 front of the vasculature (i.e., interspersed between the blood vessels and the normal cornea). In most cases this would involve perilimbal corneal injection to "protect" the

cornea from the advancing blood vessels. This method may also be utilized shortly after a corneal insult in order to prophylactically prevent corneal neovascularization. In this situation the material could be injected in the perilimbic cornea interspersed between the corneal lesion and its undesired potential limbic blood supply. Such methods may also be utilized in a similar fashion to prevent capillary invasion of transplanted corneas. In a sustained-release form injections might only be required 2-3 times per year. A steroid could also be added to the injection solution to reduce inflammation resulting from the injection itself.

Within another aspect of the present invention, methods are provided for treating or preventing neovascular glaucoma, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eye, such that the formation of blood vessels is inhibited. In one embodiment, the compound may be administered topically to the eye in order to treat or prevent early forms of neovascular glaucoma. Within other embodiments, the compound may be implanted by injection into the region of the anterior chamber angle. Within other embodiments, the compound may also be placed in any location such that the compound is continuously released into the aqueous humor. Within another aspect of the present invention, methods are provided for treating or preventing proliferative diabetic retinopathy, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eyes, such that the formation of blood vessels is inhibited.

Within particularly preferred embodiments of the invention, proliferative diabetic retinopathy may be treated by injection into the aqueous humor or the vitreous, in order to increase the local concentration of the polynucleotide, polypeptide, antagonist and/or agonist in the retina. Preferably, this treatment should

be initiated prior to the acquisition of severe disease requiring photocoagulation.

Within another aspect of the present invention, methods are provided for treating or preventing retrolental fibroplasia, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist  
5 and/or agonist to the eye, such that the formation of blood vessels is inhibited. The compound may be administered topically, via intravitreal injection and/or via intraocular implants.

Additionally, diseases, disorders, and/or conditions which can be treated, prevented, and/or diagnosed with the polynucleotides, polypeptides, agonists and/or  
10 agonists include, but are not limited to, hemangioma, arthritis, psoriasis, angiofibroma, atherosclerotic plaques, delayed wound healing, granulations, hemophilic joints, hypertrophic scars, nonunion fractures, Osler-Weber syndrome, pyogenic granuloma, scleroderma, trachoma, and vascular adhesions.

Moreover, diseases, disorders, and/or conditions and/or states, which can be  
15 treated, prevented, and/or diagnosed with the the polynucleotides, polypeptides, agonists and/or agonists include, but are not limited to, solid tumors, blood born tumors such as leukemias, tumor metastasis, Kaposi's sarcoma, benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases, for example,  
20 diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, and uveitis, delayed wound healing, endometriosis, vasculogenesis, granulations, hypertrophic scars (keloids), nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals,  
25 arteriovenous malformations, ischemic limb angiogenesis, Osler-Webber Syndrome, plaque neovascularization, telangiectasia, hemophiliac joints, angiofibroma

fibromuscular dysplasia, wound granulation, Crohn's disease, atherosclerosis, birth control agent by preventing vascularization required for embryo implantation controlling menstruation, diseases that have angiogenesis as a pathologic consequence such as cat scratch disease (Rochele minalia quintosa), ulcers (Helicobacter pylori),

5   Bartonellosis and bacillary angiomatosis.

          In one aspect of the birth control method, an amount of the compound sufficient to block embryo implantation is administered before or after intercourse and fertilization have occurred, thus providing an effective method of birth control, possibly a "morning after" method. Polynucleotides, polypeptides, agonists and/or

10   agonists may also be used in controlling menstruation or administered as either a peritoneal lavage fluid or for peritoneal implantation in the treatment of endometriosis.

          Polynucleotides, polypeptides, agonists and/or agonists of the present invention may be incorporated into surgical sutures in order to prevent stitch

15   granulomas.

          Polynucleotides, polypeptides, agonists and/or agonists may be utilized in a wide variety of surgical procedures. For example, within one aspect of the present invention a compositions (in the form of, for example, a spray or film) may be utilized to coat or spray an area prior to removal of a tumor, in order to isolate normal

20   surrounding tissues from malignant tissue, and/or to prevent the spread of disease to surrounding tissues. Within other aspects of the present invention, compositions (e.g., in the form of a spray) may be delivered via endoscopic procedures in order to coat tumors, or inhibit angiogenesis in a desired locale. Within yet other aspects of the present invention, surgical meshes which have been coated with anti- angiogenic

25   compositions of the present invention may be utilized in any procedure wherein a surgical mesh might be utilized. For example, within one embodiment of the



invention a surgical mesh laden with an anti-angiogenic composition may be utilized during abdominal cancer resection surgery (e.g., subsequent to colon resection) in order to provide support to the structure, and to release an amount of the anti-angiogenic factor.

5           Within further aspects of the present invention, methods are provided for treating tumor excision sites, comprising administering a polynucleotide, polypeptide, agonist and/or antagonist to the resection margins of a tumor subsequent to excision, such that the local recurrence of cancer and the formation of new blood vessels at the site is inhibited. Within one embodiment of the invention, the anti-angiogenic  
10   compound is administered directly to the tumor excision site (e.g., applied by swabbing, brushing or otherwise coating the resection margins of the tumor with the anti-angiogenic compound). Alternatively, the anti-angiogenic compounds may be incorporated into known surgical pastes prior to administration. Within particularly preferred embodiments of the invention, the anti-angiogenic compounds are applied  
15   after hepatic resections for malignancy, and after neurosurgical operations.

          Within one aspect of the present invention, polynucleotides, polypeptides, agonists and/or antagonists may be administered to the resection margin of a wide variety of tumors, including for example, breast, colon, brain and hepatic tumors. For example, within one embodiment of the invention, anti-angiogenic compounds may  
20   be administered to the site of a neurological tumor subsequent to excision, such that the formation of new blood vessels at the site are inhibited.

          The polynucleotides, polypeptides, agonists and/or antagonists of the present invention may also be administered along with other anti-angiogenic factors.

Representative examples of other anti-angiogenic factors include: Anti-Invasive  
25   Factor, retinoic acid and derivatives thereof, paclitaxel, Suramin, Tissue Inhibitor of Metalloproteinase-1, Tissue Inhibitor of Metalloproteinase-2, Plasminogen Activator

Inhibitor-1, Plasminogen Activator Inhibitor-2, and various forms of the lighter "d group" transition metals.

Lighter "d group" transition metals include, for example, vanadium, molybdenum, tungsten, titanium, niobium, and tantalum species. Such transition metal species may form transition metal complexes. Suitable complexes of the above-mentioned transition metal species include oxo transition metal complexes.

Representative examples of vanadium complexes include oxo vanadium complexes such as vanadate and vanadyl complexes. Suitable vanadate complexes include metavanadate and orthovanadate complexes such as, for example, ammonium metavanadate, sodium metavanadate, and sodium orthovanadate. Suitable vanadyl complexes include, for example, vanadyl acetylacetonate and vanadyl sulfate including vanadyl sulfate hydrates such as vanadyl sulfate mono- and trihydrates.

Representative examples of tungsten and molybdenum complexes also include oxo complexes. Suitable oxo tungsten complexes include tungstate and tungsten oxide complexes. Suitable tungstate complexes include ammonium tungstate, calcium tungstate, sodium tungstate dihydrate, and tungstic acid. Suitable tungsten oxides include tungsten (IV) oxide and tungsten (VI) oxide. Suitable oxo molybdenum complexes include molybdate, molybdenum oxide, and molybdenyl complexes. Suitable molybdate complexes include ammonium molybdate and its hydrates, sodium molybdate and its hydrates, and potassium molybdate and its hydrates. Suitable molybdenum oxides include molybdenum (VI) oxide, molybdenum (VI) oxide, and molybdic acid. Suitable molybdenyl complexes include, for example, molybdenyl acetylacetonate. Other suitable tungsten and molybdenum complexes include hydroxo derivatives derived from, for example, glycerol, tartaric acid, and sugars.

A wide variety of other anti-angiogenic factors may also be utilized within the context of the present invention. Representative examples include platelet factor 4; protamine sulphate; sulphated chitin derivatives (prepared from queen crab shells), (Murata et al., Cancer Res. 51:22-26, 1991); Sulphated Polysaccharide Peptidoglycan  
5 Complex (SP- PG) (the function of this compound may be enhanced by the presence of steroids such as estrogen, and tamoxifen citrate); Staurosporine; modulators of matrix metabolism, including for example, proline analogs, cishydroxyproline, d,L-3,4-dehydroproline, Thiaproline, alpha,alpha-dipyridyl, aminopropionitrile fumarate; 4-propyl-5-(4-pyridinyl)-2(3H)-oxazolone; Methotrexate; Mitoxantrone; Heparin;  
10 Interferons; 2 Macroglobulin-serum; ChIMP-3 (Pavloff et al., J. Bio. Chem. 267:17321-17326, 1992); Chymostatin (Tomkinson et al., Biochem J. 286:475-480, 1992); Cyclodextrin Tetradasulfate; Eponemycin; Camptothecin; Fumagillin (Ingber et al., Nature 348:555-557, 1990); Gold Sodium Thiomalate ("GST"; Matsubara and Ziff, J. Clin. Invest. 79:1440-1446, 1987); anticollagenase-serum;  
15 alpha2-antiplasmin (Holmes et al., J. Biol. Chem. 262(4):1659-1664, 1987); Bisantrene (National Cancer Institute); Lobenzarit disodium (N-(2)-carboxyphenyl-4-chloroanthronilic acid disodium or "CCA"; Takeuchi et al., Agents Actions 36:312-316, 1992); Thalidomide; Angostatic steroid; AGM-1470; carboxynaminolmidazole; and metalloproteinase inhibitors such as BB94.

20

#### **Diseases at the Cellular Level**

Diseases associated with increased cell survival or the inhibition of apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides and/or antagonists or agonists of the invention, include cancers (such as  
25 follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, including, but not limited to colon cancer, cardiac tumors, pancreatic cancer,

melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, breast cancer, prostate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation, graft v. host disease, acute graft rejection, and chronic graft rejection. In preferred embodiments, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention are used to inhibit growth, progression, and/or metasis of cancers, in particular those listed above.

Additional diseases or conditions associated with increased cell survival that could be treated, prevented or diagnosed by the polynucleotides or polypeptides, or agonists or antagonists of the invention, include, but are not limited to, progression, and/or metastases of malignancies and related disorders such as leukemia (including acute leukemias (e.g., acute lymphocytic leukemia, acute myelocytic leukemia (including myeloblastic, promyelocytic, myelomonocytic, monocytic, and erythroleukemia)) and chronic leukemias (e.g., chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia)), polycythemia vera, lymphomas (e.g., Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors including, but not limited to, sarcomas and carcinomas such as fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic

cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma, and retinoblastoma.

10 Diseases associated with increased apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, include AIDS; neurodegenerative diseases, disorders, and/or conditions (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration and brain tumor or

15 prior associated disease); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) myelodysplastic syndromes (such as aplastic anemia), graft v. host disease, ischemic injury (such as

20 that caused by myocardial infarction, stroke and reperfusion injury), liver injury (e.g., hepatitis related liver injury, ischemia/reperfusion injury, cholestosis (bile duct injury) and liver cancer); toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia.

### **Wound Healing and Epithelial Cell Proliferation**

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, for therapeutic purposes, for example, to stimulate  
5 epithelial cell proliferation and basal keratinocytes for the purpose of wound healing, and to stimulate hair follicle production and healing of dermal wounds.

Polynucleotides or polypeptides, as well as agonists or antagonists of the invention, may be clinically useful in stimulating wound healing including surgical wounds, excisional wounds, deep wounds involving damage of the dermis and epidermis, eye  
10 tissue wounds, dental tissue wounds, oral cavity wounds, diabetic ulcers, dermal ulcers, cubitus ulcers, arterial ulcers, venous stasis ulcers, burns resulting from heat exposure or chemicals, and other abnormal wound healing conditions such as uremia, malnutrition, vitamin deficiencies and complications associated with systemic treatment with steroids, radiation therapy and antineoplastic drugs and  
15 antimetabolites. Polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote dermal reestablishment subsequent to dermal loss

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to increase the adherence of skin grafts to a wound bed and  
20 to stimulate re-epithelialization from the wound bed. The following are a non-exhaustive list of grafts that polynucleotides or polypeptides, agonists or antagonists of the invention, could be used to increase adherence to a wound bed: autografts, artificial skin, allografts, autodermic graft, autoepidermic grafts, avacular grafts, Blair-Brown grafts, bone graft, brephoplastic grafts, cutis graft, delayed graft, dermic graft,  
25 epidermic graft, fascia graft, full thickness graft, heterologous graft, xenograft, homologous graft, hyperplastic graft, lamellar graft, mesh graft, mucosal graft, Ollier-

Thiersch graft, omentopapular graft, patch graft, pedicle graft, penetrating graft, split skin graft, thick split graft. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, can be used to promote skin strength and to improve the appearance of aged skin.

5           It is believed that the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, will also produce changes in hepatocyte proliferation, and epithelial cell proliferation in the lung, breast, pancreas, stomach, small intestine, and large intestine. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could promote proliferation of epithelial cells such as  
10   sebocytes, hair follicles, hepatocytes, type II pneumocytes, mucin-producing goblet cells, and other epithelial cells and their progenitors contained within the skin, lung, liver, and gastrointestinal tract. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may promote proliferation of endothelial cells, keratinocytes, and basal keratinocytes.

15           The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to reduce the side effects of gut toxicity that result from radiation, chemotherapy treatments or viral infections. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may have a cytoprotective effect on the small intestine mucosa. The polynucleotides or  
20   polypeptides, and/or agonists or antagonists of the invention, may also stimulate healing of mucositis (mouth ulcers) that result from chemotherapy and viral infections.

          The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could further be used in full regeneration of skin in full and partial  
25   thickness skin defects, including burns, (i.e., repopulation of hair follicles, sweat glands, and sebaceous glands), treatment of other skin defects such as psoriasis. The

polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat epidermolysis bullosa, a defect in adherence of the epidermis to the underlying dermis which results in frequent, open and painful blisters by accelerating reepithelialization of these lesions. The polynucleotides or polypeptides, and/or

5 agonists or antagonists of the invention, could also be used to treat gastric and duodenal ulcers and help heal by scar formation of the mucosal lining and regeneration of glandular mucosa and duodenal mucosal lining more rapidly. Inflammatory bowel diseases, such as Crohn's disease and ulcerative colitis, are diseases which result in destruction of the mucosal surface of the small or large

10 intestine, respectively. Thus, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote the resurfacing of the mucosal surface to aid more rapid healing and to prevent progression of inflammatory bowel disease. Treatment with the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, is expected to have a significant effect on the production

15 of mucus throughout the gastrointestinal tract and could be used to protect the intestinal mucosa from injurious substances that are ingested or following surgery. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat diseases associated with the under expression of the polynucleotides of the invention.

20 Moreover, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to prevent and heal damage to the lungs due to various pathological states. A growth factor such as the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, which could stimulate proliferation and differentiation and promote the repair of alveoli and bronchiolar epithelium to

25 prevent or treat acute or chronic lung damage. For example, emphysema, which results in the progressive loss of alveoli, and inhalation injuries, i.e., resulting from



smoke inhalation and burns, that cause necrosis of the bronchiolar epithelium and alveoli could be effectively treated, prevented, and/or diagnosed using the polynucleotides or polypeptides, and/or agonists or antagonists of the invention.

Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to stimulate the proliferation of and differentiation of type II pneumocytes, which may help treat or prevent disease such as hyaline membrane diseases, such as infant respiratory distress syndrome and bronchopulmonary dysplasia, in premature infants.

The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could stimulate the proliferation and differentiation of hepatocytes and, thus, could be used to alleviate or treat liver diseases and pathologies such as fulminant liver failure caused by cirrhosis, liver damage caused by viral hepatitis and toxic substances (i.e., acetaminophen, carbon tetrachloride and other hepatotoxins known in the art).

In addition, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat or prevent the onset of diabetes mellitus. In patients with newly diagnosed Types I and II diabetes, where some islet cell function remains, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to maintain the islet function so as to alleviate, delay or prevent permanent manifestation of the disease. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used as an auxiliary in islet cell transplantation to improve or promote islet cell function.

## **Neurological Diseases**

Nervous system diseases, disorders, and/or conditions, which can be treated, prevented, and/or diagnosed with the compositions of the invention (e.g., polypeptides, polynucleotides, and/or agonists or antagonists), include, but are not limited to, nervous system injuries, and diseases, disorders, and/or conditions which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated, prevented, and/or diagnosed in a patient (including human and non-human mammalian patients) according to the invention, include but are not limited to, the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems: (1) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia; (2) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries; (3) malignant lesions, in which a portion of the nervous system is destroyed or injured by malignant tissue which is either a nervous system associated malignancy or a malignancy derived from non-nervous system tissue; (4) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis; (5) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis (ALS); (6) lesions associated with nutritional diseases, disorders, and/or conditions, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic

acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration; (7) neurological lesions associated with systemic diseases including, but not limited to, diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis; (8) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and (9) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including, but not limited to, multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

In a preferred embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to protect neural cells from the damaging effects of cerebral hypoxia. According to this embodiment, the compositions of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral hypoxia. In one aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral ischemia. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral infarction. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose or prevent neural cell injury associated with a stroke. In a further aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with a heart attack.

The compositions of the invention which are useful for treating or preventing a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, compositions of the invention which elicit any of the following effects may be useful according to the invention: (1) increased survival time of neurons in culture; (2) increased sprouting of neurons in culture or *in vivo*; (3) increased production of a neuron-associated molecule in culture or *in vivo*, e.g., choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or (4) decreased symptoms of neuron dysfunction *in vivo*. Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may routinely be measured using a method set forth herein or otherwise known in the art, such as, for example, the method set forth in Arakawa et al. (J. Neurosci. 10:3507-3515 (1990)); increased sprouting of neurons may be detected by methods known in the art, such as, for example, the methods set forth in Pestronk et al. (Exp. Neurol. 70:65-82 (1980)) or Brown et al. (Ann. Rev. Neurosci. 4:17-42 (1981)); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., using techniques known in the art and depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron diseases, disorders, and/or conditions that may be treated, prevented, and/or diagnosed according to the invention include, but are not limited to, diseases, disorders, and/or conditions such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous

system, as well as diseases, disorders, and/or conditions that selectively affect neurons such as amyotrophic lateral sclerosis, and including, but not limited to, progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary  
5   Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

### **Infectious Disease**

A polypeptide or polynucleotide and/or agonist or antagonist of the present  
10   invention can be used to treat, prevent, and/or diagnose infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated, prevented, and/or diagnosed. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively,  
15   polypeptide or polynucleotide and/or agonist or antagonist of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated, prevented, and/or diagnosed by a polynucleotide or  
20   polypeptide and/or agonist or antagonist of the present invention. Examples of viruses, include, but are not limited to Examples of viruses, include, but are not limited to the following DNA and RNA viruses and viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Dengue, EBV, HIV, Flaviviridae, Hepadnaviridae  
25   (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae),

Orthomyxoviridae (e.g., Influenza A, Influenza B, and parainfluenza), Papiloma virus, Papovaviridae, Parvoviridae, Picornaviridae, Poxviridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a  
5 variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, respiratory syncytial virus, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), Japanese B encephalitis, Junin, Chikungunya, Rift Valley fever, yellow fever, meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox,  
10 hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or  
15 antagonists of the invention are used to treat, prevent, and/or diagnose: meningitis, Dengue, EBV, and/or hepatitis (e.g., hepatitis B). In an additional specific embodiment polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat patients nonresponsive to one or more other commercially available hepatitis vaccines. In a further specific embodiment polynucleotides, polypeptides, or  
20 agonists or antagonists of the invention are used to treat, prevent, and/or diagnose AIDS.

Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to,  
25 include, but not limited to, the following Gram-Negative and Gram-positive bacteria and bacterial families and fungi: Actinomycetales (e.g., Corynebacterium,

Mycobacterium, Norcardia), Cryptococcus neoformans, Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia (e.g., Borrelia burgdorferi), Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, E. coli (e.g., Enterotoxigenic E. coli and Enterohemorrhagic E. coli), Enterobacteriaceae (Klebsiella, Salmonella (e.g., Salmonella typhi, and Salmonella paratyphi), Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, Mycobacterium leprae, Vibrio cholerae, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Meningococcal), Meisseria meningitidis, Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus (e.g., Heamophilus influenza type B), Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, Shigella spp., Staphylococcal, Meningiococcal, Pneumococcal and Streptococcal (e.g., Streptococcus pneumoniae and Group B Streptococcus). These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to:

bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis (e.g., meningitis types A and B), Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections.

Polynucleotides or polypeptides, agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, agonists or antagonists of the invention

are used to treat, prevent, and/or diagnose: tetanus, Diphtheria, botulism, and/or meningitis type B.

Moreover, parasitic agents causing disease or symptoms that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, the following families or class: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, 5 Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas and Sporozoans (e.g., Plasmodium virax, Plasmodium falciparum, Plasmodium malariae and Plasmodium 10 ovale). These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), malaria, pregnancy complications, and toxoplasmosis. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, 15 and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose malaria.

Preferably, treatment or prevention using a polypeptide or polynucleotide and/or agonist or antagonist of the present invention could either be by administering 20 an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

25

### **Regeneration**



A polynucleotide or polypeptide and/or agonist or antagonist of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, 5 trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteoarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal 10 or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

Moreover, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase regeneration of tissues difficult to heal. For example, 15 increased tendon/ligament regeneration would quicken recovery time after damage.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated, prevented, and/or diagnosed include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue 20 regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide and/or agonist or antagonist of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated, prevented, 25 and/or diagnosed using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic diseases, disorders, and/or

conditions (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stroke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, 5 Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated, prevented, and/or diagnosed using the polynucleotide or polypeptide and/or agonist or antagonist of the present invention.

### **Chemotaxis**

10 A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may have chemotaxis activity. A chemotactic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then 15 fight off and/or heal the particular trauma or abnormality.

A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase chemotactic activity of particular cells. These chemotactic molecules can then be used to treat, prevent, and/or diagnose inflammation, infection, hyperproliferative diseases, disorders, and/or conditions, or any immune system 20 disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotactic molecules can be used to treat, prevent, and/or diagnose wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat, prevent, and/or diagnose wounds.

25 It is also contemplated that a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may inhibit chemotactic activity. These molecules

could also be used to treat, prevent, and/or diagnose diseases, disorders, and/or conditions. Thus, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention could be used as an inhibitor of chemotaxis.

5

### **Binding Activity**

A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit  
10 (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural  
15 or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

20 Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially  
25 containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

5           Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a  
10       standard.

          Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

15           Additionally, the receptor to which a polypeptide of the invention binds can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting (Coligan, et al., Current Protocols in Immun., 1(2), Chapter 5, (1991)). For example, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the polypeptides, for  
20       example, NIH3T3 cells which are known to contain multiple receptors for the FGF family proteins, and SC-3 cells, and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the polypeptides. Transfected cells which are grown on glass slides are exposed to the polypeptide of the present invention, after they have been labelled. The polypeptides  
25       can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase.

Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an iterative sub-pooling and re-screening process, eventually yielding a single clones that encodes the putative receptor.

5           As an alternative approach for receptor identification, the labeled polypeptides can be photoaffinity linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE analysis and exposed to X-ray film. The labeled complex containing the receptors of the polypeptides can be excised, resolved into peptide fragments, and subjected to protein  
10   microsequencing. The amino acid sequence obtained from microsequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the genes encoding the putative receptors.

Moreover, the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling") may be  
15   employed to modulate the activities of polypeptides of the invention thereby effectively generating agonists and antagonists of polypeptides of the invention. See generally, U.S. Patent Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, and 5,837,458, and Patten, P. A., et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, S. Trends Biotechnol. 16(2):76-82 (1998); Hansson, L. O., et al., J. Mol.  
20   Biol. 287:265-76 (1999); and Lorenzo, M. M. and Blasco, R. Biotechniques 24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference). In one embodiment, alteration of polynucleotides and corresponding polypeptides of the invention may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments into a desired polynucleotide  
25   sequence of the invention molecule by homologous, or site-specific, recombination. In another embodiment, polynucleotides and corresponding polypeptides of the

invention may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of the polypeptides of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules. In preferred embodiments, the heterologous molecules are family members. In further preferred embodiments, the heterologous molecule is a growth factor such as, for example, platelet-derived growth factor (PDGF), insulin-like growth factor (IGF-I), transforming growth factor (TGF)-alpha, epidermal growth factor (EGF), fibroblast growth factor (FGF), TGF-beta, bone morphogenetic protein (BMP)-2, BMP-4, BMP-5, BMP-6, BMP-7, activins A and B, decapentaplegic(dpp), 60A, OP-2, dorsalin, growth differentiation factors (GDFs), nodal, MIS, inhibin-alpha, TGF-beta1, TGF-beta2, TGF-beta3, TGF-beta5, and glial-derived neurotrophic factor (GDNF).

Other preferred fragments are biologically active fragments of the polypeptides of the invention. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

Additionally, this invention provides a method of screening compounds to identify those which modulate the action of the polypeptide of the present invention. An example of such an assay comprises combining a mammalian fibroblast cell, a the polypeptide of the present invention, the compound to be screened and 3[H] thymidine under cell culture conditions where the fibroblast cell would normally proliferate. A control assay may be performed in the absence of the compound to be screened and compared to the amount of fibroblast proliferation in the presence of the

compound to determine if the compound stimulates proliferation by determining the uptake of 3[H] thymidine in each case. The amount of fibroblast cell proliferation is measured by liquid scintillation chromatography which measures the incorporation of 3[H] thymidine. Both agonist and antagonist compounds may be identified by this  
5 procedure.

In another method, a mammalian cell or membrane preparation expressing a receptor for a polypeptide of the present invention is incubated with a labeled polypeptide of the present invention in the presence of the compound. The ability of the compound to enhance or block this interaction could then be measured.  
10 Alternatively, the response of a known second messenger system following interaction of a compound to be screened and the receptor is measured and the ability of the compound to bind to the receptor and elicit a second messenger response is measured to determine if the compound is a potential agonist or antagonist. Such second messenger systems include but are not limited to, cAMP guanylate cyclase,  
15 ion channels or phosphoinositide hydrolysis.

All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat, prevent, and/or diagnose disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays  
20 can discover agents which may inhibit or enhance the production of the polypeptides of the invention from suitably manipulated cells or tissues. Therefore, the invention includes a method of identifying compounds which bind to the polypeptides of the invention comprising the steps of: (a) incubating a candidate binding compound with the polypeptide; and (b) determining if binding has occurred. Moreover, the  
25 invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with the polypeptide, (b) assaying a

biological activity , and (b) determining if a biological activity of the polypeptide has been altered.

Also, one could identify molecules bind a polypeptide of the invention experimentally by using the beta-pleated sheet regions contained in the polypeptide sequence of the protein. Accordingly, specific embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, the amino acid sequence of each beta pleated sheet regions in a disclosed polypeptide sequence. Additional embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, any combination or all of contained in the polypeptide sequences of the invention. Additional preferred embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, the amino acid sequence of each of the beta pleated sheet regions in one of the polypeptide sequences of the invention. Additional embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, any combination or all of the beta pleated sheet regions in one of the polypeptide sequences of the invention.

### **Targeted Delivery**

In another embodiment, the invention provides a method of delivering compositions to targeted cells expressing a receptor for a polypeptide of the invention, or cells expressing a cell bound form of a polypeptide of the invention.

As discussed herein, polypeptides or antibodies of the invention may be associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic and/or covalent interactions. In one embodiment, the invention provides a method for the specific delivery of compositions of the invention to cells by administering polypeptides of the invention



(including antibodies) that are associated with heterologous polypeptides or nucleic acids. In one example, the invention provides a method for delivering a therapeutic protein into the targeted cell. In another example, the invention provides a method for delivering a single stranded nucleic acid (e.g., antisense or ribozymes) or double  
5 stranded nucleic acid (e.g., DNA that can integrate into the cell's genome or replicate episomally and that can be transcribed) into the targeted cell.

In another embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides of the invention (e.g., polypeptides of the invention or antibodies of the invention) in  
10 association with toxins or cytotoxic prodrugs.

By "toxin" is meant compounds that bind and activate endogenous cytotoxic effector systems, radioisotopes, holotoxins, modified toxins, catalytic subunits of toxins, or any molecules or enzymes not normally present in or on the surface of a cell that under defined conditions cause the cell's death. Toxins that may be used  
15 according to the methods of the invention include, but are not limited to, radioisotopes known in the art, compounds such as, for example, antibodies (or complement fixing containing portions thereof) that bind an inherent or induced endogenous cytotoxic effector system, thymidine kinase, endonuclease, RNase, alpha toxin, ricin, abrin, *Pseudomonas* exotoxin A, diphtheria toxin, saporin, momordin, gelonin, pokeweed  
20 antiviral protein, alpha-sarcin and cholera toxin. By "cytotoxic prodrug" is meant a non-toxic compound that is converted by an enzyme, normally present in the cell, into a cytotoxic compound. Cytotoxic prodrugs that may be used according to the methods of the invention include, but are not limited to, glutamyl derivatives of benzoic acid mustard alkylating agent, phosphate derivatives of etoposide or  
25 mitomycin C, cytosine arabinoside, daunorubisin, and phenoxyacetamide derivatives of doxorubicin.

**Drug Screening**

Further contemplated is the use of the polypeptides of the present invention, or the polynucleotides encoding these polypeptides, to screen for molecules which  
5 modify the activities of the polypeptides of the present invention. Such a method would include contacting the polypeptide of the present invention with a selected compound(s) suspected of having antagonist or agonist activity, and assaying the activity of these polypeptides following binding.

This invention is particularly useful for screening therapeutic compounds by  
10 using the polypeptides of the present invention, or binding fragments thereof, in any of a variety of drug screening techniques. The polypeptide or fragment employed in such a test may be affixed to a solid support, expressed on a cell surface, free in solution, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids  
15 expressing the polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. One may measure, for example, the formulation of complexes between the agent being tested and a polypeptide of the present invention.

Thus, the present invention provides methods of screening for drugs or any  
20 other agents which affect activities mediated by the polypeptides of the present invention. These methods comprise contacting such an agent with a polypeptide of the present invention or a fragment thereof and assaying for the presence of a complex between the agent and the polypeptide or a fragment thereof, by methods well known in the art. In such a competitive binding assay, the agents to screen are typically  
25 labeled. Following incubation, free agent is separated from that present in bound

form, and the amount of free or uncomplexed label is a measure of the ability of a particular agent to bind to the polypeptides of the present invention.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the polypeptides of the present invention, and is described in great detail in European Patent Application 84/03564, published on September 13, 1984, which is incorporated herein by reference herein. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with polypeptides of the present invention and washed. Bound polypeptides are then detected by methods well known in the art. Purified polypeptides are coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies may be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding polypeptides of the present invention specifically compete with a test compound for binding to the polypeptides or fragments thereof. In this manner, the antibodies are used to detect the presence of any peptide which shares one or more antigenic epitopes with a polypeptide of the invention.

20

#### Antisense And Ribozyme (Antagonists)

In specific embodiments, antagonists according to the present invention are nucleic acids corresponding to the sequences contained in SEQ ID NO:X, or the complementary strand thereof, and/or to nucleotide sequences contained a deposited clone. In one embodiment, antisense sequence is generated internally by the organism, in another embodiment, the antisense sequence is separately administered

25

(see, for example, O'Connor, *Neurochem.*, 56:560 (1991). *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988).

Antisense technology can be used to control gene expression through antisense DNA or RNA, or through triple-helix formation. Antisense techniques are discussed for  
5 example, in Okano, *Neurochem.*, 56:560 (1991); *Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression*, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance, Lee et al., *Nucleic Acids Research*, 6:3073 (1979); Cooney et al., *Science*, 241:456 (1988); and Dervan et al., *Science*, 251:1300 (1991). The methods are based on binding of a polynucleotide to a complementary  
10 DNA or RNA.

For example, the use of c-myc and c-myb antisense RNA constructs to inhibit the growth of the non-lymphocytic leukemia cell line HL-60 and other cell lines was previously described. (Wickstrom et al. (1988); Anfossi et al. (1989)). These experiments were performed in vitro by incubating cells with the oligoribonucleotide.  
15 A similar procedure for in vivo use is described in WO 91/15580. Briefly, a pair of oligonucleotides for a given antisense RNA is produced as follows: A sequence complementary to the first 15 bases of the open reading frame is flanked by an EcoRI site on the 5' end and a HindIII site on the 3' end. Next, the pair of oligonucleotides is heated at 90°C for one minute and then annealed in 2X ligation buffer (20mM TRIS  
20 HCl pH 7.5, 10mM MgCl<sub>2</sub>, 10mM dithiothreitol (DTT) and 0.2 mM ATP) and then ligated to the EcoRI/Hind III site of the retroviral vector PMV7 (WO 91/15580).

For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide  
25 is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of the receptor. The antisense

RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into receptor polypeptide.

In one embodiment, the antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector or  
5 a portion thereof, is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the antisense nucleic acid of the invention. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard  
10 in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding a polypeptide of the invention, or fragments thereof, can be by any promoter known in the art to act in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early  
15 promoter region (Bernoist and Chambon, Nature, 29:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., Cell, 22:787-797 (1980), the herpes thymidine promoter (Wagner et al., Proc. Natl. Acad. Sci. U.S.A., 78:1441-1445 (1981), the regulatory sequences of the metallothionein gene (Brinster et al., Nature, 296:39-42 (1982)), etc.

20 The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a gene of interest. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA,  
25 forming a stable duplex; in the case of double stranded antisense nucleic acids of the invention, a single strand of the duplex DNA may thus be tested, or triplex formation

may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the larger the hybridizing nucleic acid, the more base mismatches with a RNA sequence of the invention it may contain and still form a stable duplex (or triplex as the case may be).

- 5 One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, *e.g.*, the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., *Nature*, 372:333-335 (1994). Thus, oligonucleotides complementary to either the 5' - or 3' - non- translated, non-coding regions of a polynucleotide sequence of the invention could be used in an antisense approach to inhibit translation of endogenous mRNA.

15 Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5' -, 3' - or coding region of mRNA, antisense nucleic acids should be at least six

20 nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

The polynucleotides of the invention can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-

25 stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization,

etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci., 84:648-652 (1987); PCT Publication NO: WO88/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication NO: WO89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al., BioTechniques, 6:958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res., 5:539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least  
5 one modified phosphate backbone selected from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an a-anomeric  
10 oligonucleotide. An a-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run parallel to each other (Gautier et al., Nucl. Acids Res., 15:6625-6641 (1987)). The oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al., Nucl. Acids Res., 15:6131-6148 (1987)), or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett.  
15 215:327-330 (1987)).

Polynucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al.  
20 (Nucl. Acids Res., 16:3209 (1988)), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., Proc. Natl. Acad. Sci. U.S.A., 85:7448-7451 (1988)), etc.

While antisense nucleotides complementary to the coding region sequence of the invention could be used, those complementary to the transcribed untranslated  
25 region are most preferred.



Potential antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published October 4, 1990; Sarver et al, Science, 247:1222-1225 (1990). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy mRNAs  
5 corresponding to the polynucleotides of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5' -UG-3'. The construction and production of hammerhead ribozymes is well  
10 known in the art and is described more fully in Haseloff and Gerlach, Nature, 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within each nucleotide sequence disclosed in the sequence listing. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the mRNA corresponding to the polynucleotides of the invention; i.e., to  
15 increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

As in the antisense approach, the ribozymes of the invention can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the polynucleotides of the invention in vivo.  
20 DNA constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive promoter, such as, for example, pol III or pol II promoter, so that transfected cells will produce sufficient quantities of  
25 the ribozyme to destroy endogenous messages and inhibit translation. Since

ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Antagonist/agonist compounds may be employed to inhibit the cell growth and proliferation effects of the polypeptides of the present invention on neoplastic  
5 cells and tissues, i.e. stimulation of angiogenesis of tumors, and, therefore, retard or prevent abnormal cellular growth and proliferation, for example, in tumor formation or growth.

The antagonist/agonist may also be employed to prevent hyper-vascular diseases, and prevent the proliferation of epithelial lens cells after extracapsular  
10 cataract surgery. Prevention of the mitogenic activity of the polypeptides of the present invention may also be desirable in cases such as restenosis after balloon angioplasty.

The antagonist/agonist may also be employed to prevent the growth of scar tissue during wound healing.

15 The antagonist/agonist may also be employed to treat, prevent, and/or diagnose the diseases described herein.

Thus, the invention provides a method of treating or preventing diseases, disorders, and/or conditions, including but not limited to the diseases, disorders, and/or conditions listed throughout this application, associated with  
20 overexpression of a polynucleotide of the present invention by administering to a patient (a) an antisense molecule directed to the polynucleotide of the present invention, and/or (b) a ribozyme directed to the polynucleotide of the present invention.

invention, and/or (b) a ribozyme directed to the polynucleotide of the present  
25 invention.

**Other Activities**

The polypeptide of the present invention, as a result of the ability to stimulate vascular endothelial cell growth, may be employed in treatment for stimulating re-vascularization of ischemic tissues due to various disease conditions such as

5 thrombosis, arteriosclerosis, and other cardiovascular conditions. These polypeptide may also be employed to stimulate angiogenesis and limb regeneration, as discussed above.

The polypeptide may also be employed for treating wounds due to injuries, burns, post-operative tissue repair, and ulcers since they are mitogenic to various cells

10 of different origins, such as fibroblast cells and skeletal muscle cells, and therefore, facilitate the repair or replacement of damaged or diseased tissue.

The polypeptide of the present invention may also be employed stimulate neuronal growth and to treat, prevent, and/or diagnose neuronal damage which occurs in certain neuronal disorders or neuro-degenerative conditions such as Alzheimer's

15 disease, Parkinson's disease, and AIDS-related complex. The polypeptide of the invention may have the ability to stimulate chondrocyte growth, therefore, they may be employed to enhance bone and periodontal regeneration and aid in tissue transplants or bone grafts.

The polypeptide of the present invention may be also be employed to prevent

20 skin aging due to sunburn by stimulating keratinocyte growth.

The polypeptide of the invention may also be employed for preventing hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth. Along the same lines, the polypeptides of the present invention may be employed to stimulate growth and differentiation of hematopoietic cells and bone

25 marrow cells when used in combination with other cytokines.

The polypeptide of the invention may also be employed to maintain organs before transplantation or for supporting cell culture of primary tissues.

The polypeptide of the present invention may also be employed for inducing tissue of mesodermal origin to differentiate in early embryos.

5

The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

10 The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, polypeptides or polynucleotides and/or agonist or antagonists of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing,  
15 utilization, and storage of energy.

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, cardiac rhythms, depression (including depressive diseases, disorders, and/or conditions), tendency for violence, tolerance for pain, reproductive  
20 capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins,  
25 minerals, cofactors or other nutritional components.

**Other Preferred Embodiments**

Other preferred embodiments of the claimed invention include an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 50 contiguous nucleotides in the nucleotide  
5 sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Clone Sequence and ending with the nucleotide at about the position of the 3'  
10 Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the Start Codon and ending with the nucleotide at about the position of the 3'  
15 Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Similarly preferred is a nucleic acid molecule wherein said sequence of contiguous nucleotides is included in the nucleotide sequence of SEQ ID NO:X in the range of positions beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the  
20 nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 150 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

Further preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least about 500 contiguous nucleotides in the nucleotide sequence of SEQ ID NO:X.

A further preferred embodiment is a nucleic acid molecule comprising a  
5 nucleotide sequence which is at least 95% identical to the nucleotide sequence of SEQ ID NO:X beginning with the nucleotide at about the position of the 5' Nucleotide of the First Amino Acid of the Signal Peptide and ending with the nucleotide at about the position of the 3' Nucleotide of the Clone Sequence as defined for SEQ ID NO:X in Table 1.

10 A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence of SEQ ID NO:X.

Also preferred is an isolated nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule, wherein said nucleic  
15 acid molecule which hybridizes does not hybridize under stringent hybridization conditions to a nucleic acid molecule having a nucleotide sequence consisting of only A residues or of only T residues.

Also preferred is a composition of matter comprising a DNA molecule which comprises a human cDNA clone identified by a cDNA Clone Identifier in Table 1,  
20 which DNA molecule is contained in the material deposited with the American Type Culture Collection and given the ATCC Deposit Number shown in Table 1 for said cDNA Clone Identifier.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous  
25 nucleotides in the nucleotide sequence of a human cDNA clone identified by a cDNA

Clone Identifier in Table 1, which DNA molecule is contained in the deposit given the ATCC Deposit Number shown in Table 1.

Also preferred is an isolated nucleic acid molecule, wherein said sequence of at least 50 contiguous nucleotides is included in the nucleotide sequence of the complete open reading frame sequence encoded by said human cDNA clone.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 150 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to sequence of at least 500 contiguous nucleotides in the nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to the complete nucleotide sequence encoded by said human cDNA clone.

A further preferred embodiment is a method for detecting in a biological sample a nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing a nucleotide sequence of at least one nucleic acid molecule in said sample with a sequence selected from said group and determining whether the sequence of said nucleic acid molecule in said sample is at least 95% identical to said selected sequence.

Also preferred is the above method wherein said step of comparing sequences comprises determining the extent of nucleic acid hybridization between nucleic acid molecules in said sample and a nucleic acid molecule comprising said sequence selected from said group. Similarly, also preferred is the above method wherein said  
5 step of comparing sequences is performed by comparing the nucleotide sequence determined from a nucleic acid molecule in said sample with said sequence selected from said group. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

A further preferred embodiment is a method for identifying the species, tissue  
10 or cell type of a biological sample which method comprises a step of detecting nucleic acid molecules in said sample, if any, comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by  
15 a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for identifying the species, tissue or cell type of a biological sample can comprise a step of detecting nucleic acid molecules comprising a  
20 nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted  
25 protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject nucleic acid molecules, if any,



comprising a nucleotide sequence that is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA  
5 Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

The method for diagnosing a pathological condition can comprise a step of detecting nucleic acid molecules comprising a nucleotide sequence in a panel of at least two nucleotide sequences, wherein at least one sequence in said panel is at least  
10 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from said group.

Also preferred is a composition of matter comprising isolated nucleic acid molecules wherein the nucleotide sequences of said nucleic acid molecules comprise a panel of at least two nucleotide sequences, wherein at least one sequence in said  
15 panel is at least 95% identical to a sequence of at least 50 contiguous nucleotides in a sequence selected from the group consisting of: a nucleotide sequence of SEQ ID NO:X wherein X is any integer as defined in Table 1; and a nucleotide sequence encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA  
20 clone in Table 1. The nucleic acid molecules can comprise DNA molecules or RNA molecules.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1.

25 Also preferred is a polypeptide, wherein said sequence of contiguous amino acids is included in the amino acid sequence of SEQ ID NO:Y in the range of

positions beginning with the residue at about the position of the First Amino Acid of the Secreted Portion and ending with the residue at about the Last Amino Acid of the Open Reading Frame as set forth for SEQ ID NO:Y in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence  
5 at least 95% identical to a sequence of at least about 30 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of SEQ ID NO:Y.

10 Further preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the complete amino acid sequence of SEQ ID NO:Y.

Further preferred is an isolated polypeptide comprising an amino acid sequence at least 90% identical to a sequence of at least about 10 contiguous amino  
15 acids in the complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is a polypeptide wherein said sequence of contiguous amino acids is included in the amino acid sequence of a secreted portion of the secreted  
20 protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 30 contiguous amino acids in the  
25 amino acid sequence of the secreted portion of the protein encoded by a human cDNA

clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence of at least about 100 contiguous amino acids in the amino acid sequence of the secreted portion of the protein encoded by a human  
5 cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is an isolated polypeptide comprising an amino acid sequence at least 95% identical to the amino acid sequence of the secreted portion of the protein  
10 encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Further preferred is an isolated antibody which binds specifically to a polypeptide comprising an amino acid sequence that is at least 90% identical to a  
15 sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table  
20 1.

Further preferred is a method for detecting in a biological sample a polypeptide comprising an amino acid sequence which is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as  
25 defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained

in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1; which method comprises a step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group and determining whether the sequence of said polypeptide molecule in said sample is  
5 at least 90% identical to said sequence of at least 10 contiguous amino acids.

Also preferred is the above method wherein said step of comparing an amino acid sequence of at least one polypeptide molecule in said sample with a sequence selected from said group comprises determining the extent of specific binding of polypeptides in said sample to an antibody which binds specifically to a polypeptide  
10 comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit  
15 with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method wherein said step of comparing sequences is performed by comparing the amino acid sequence determined from a polypeptide molecule in said sample with said sequence selected from said group.

Also preferred is a method for identifying the species, tissue or cell type of a  
20 biological sample which method comprises a step of detecting polypeptide molecules in said sample, if any, comprising an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted  
25 protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in

Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

Also preferred is the above method for identifying the species, tissue or cell type of a biological sample, which method comprises a step of detecting polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the above group.

Also preferred is a method for diagnosing in a subject a pathological condition associated with abnormal structure or expression of a gene encoding a secreted protein identified in Table 1, which method comprises a step of detecting in a biological sample obtained from said subject polypeptide molecules comprising an amino acid sequence in a panel of at least two amino acid sequences, wherein at least one sequence in said panel is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

In any of these methods, the step of detecting said polypeptide molecules includes using an antibody.

Also preferred is an isolated nucleic acid molecule comprising a nucleotide sequence which is at least 95% identical to a nucleotide sequence encoding a polypeptide wherein said polypeptide comprises an amino acid sequence that is at least 90% identical to a sequence of at least 10 contiguous amino acids in a sequence selected from the group consisting of: an amino acid sequence of SEQ ID NO:Y

wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

5           Also preferred is an isolated nucleic acid molecule, wherein said nucleotide sequence encoding a polypeptide has been optimized for expression of said polypeptide in a prokaryotic host.

          Also preferred is an isolated nucleic acid molecule, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of: an amino  
10   acid sequence of SEQ ID NO:Y wherein Y is any integer as defined in Table 1; and a complete amino acid sequence of a secreted protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA clone in Table 1.

          Further preferred is a method of making a recombinant vector comprising  
15   inserting any of the above isolated nucleic acid molecule into a vector. Also preferred is the recombinant vector produced by this method. Also preferred is a method of making a recombinant host cell comprising introducing the vector into a host cell, as well as the recombinant host cell produced by this method.

          Also preferred is a method of making an isolated polypeptide comprising  
20   culturing this recombinant host cell under conditions such that said polypeptide is expressed and recovering said polypeptide. Also preferred is this method of making an isolated polypeptide, wherein said recombinant host cell is a eukaryotic cell and said polypeptide is a secreted portion of a human secreted protein comprising an amino acid sequence selected from the group consisting of: an amino acid sequence of  
25   SEQ ID NO:Y beginning with the residue at the position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y wherein Y is an integer set forth in Table 1 and

said position of the First Amino Acid of the Secreted Portion of SEQ ID NO:Y is defined in Table 1; and an amino acid sequence of a secreted portion of a protein encoded by a human cDNA clone identified by a cDNA Clone Identifier in Table 1 and contained in the deposit with the ATCC Deposit Number shown for said cDNA  
5 clone in Table 1. The isolated polypeptide produced by this method is also preferred.

Also preferred is a method of treatment of an individual in need of an increased level of a secreted protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to  
10 increase the level of said protein activity in said individual.

The above-recited applications have uses in a wide variety of hosts. Such hosts include, but are not limited to, human, murine, rabbit, goat, guinea pig, camel, horse, mouse, rat, hamster, pig, micro-pig, chicken, goat, cow, sheep, dog, cat, non-human primate, and human. In specific embodiments, the host is a mouse, rabbit,  
15 goat, guinea pig, chicken, rat, hamster, pig, sheep, dog or cat. In preferred embodiments, the host is a mammal. In most preferred embodiments, the host is a human.

In specific embodiments of the invention, for each "Contig ID" listed in the fourth column of Table 2, preferably excluded are one or more polynucleotides  
20 comprising, or alternatively consisting of, a nucleotide sequence referenced in the fifth column of Table 2 and described by the general formula of a-b, whereas a and b are uniquely determined for the corresponding SEQ ID NO:X referred to in column 3 of Table 2. Further specific embodiments are directed to polynucleotide sequences excluding one, two, three, four, or more of the specific polynucleotide sequences  
25 referred to in the fifth column of Table 2. In no way is this listing meant to encompass all of the sequences which may be excluded by the general formula, it is just a

representative example. All references available through these accessions are hereby incorporated by reference in their entirety.

TABLE 2

Gene No.	cDNA Clone ID	NT SEQ ID NO: X	Contig ID	Public Accession Numbers
7	HPIAA80	17	829972	H22229, H25153, AA542856, C02428
10	HTWBZ57	20	828030	T77691, R35906, H48028, R89713, H53573, H53894, H60420, H60465, H80671, H81664, H82924, H82925, H88095, H88096, N20861, N24045, N27703, N35866, N74576, N80028, W06990, W51755, AA084697, AA215459, AA428087, AA428132, AA281739, AA281896, AA527742, AA534945, AA610391, AA639542, AA749283, AA862628, AA878674, AA912335, AA995185
11	HCFBV39	21	830525	R65909, N20260, N29164, AA012866, AA147481, AA147881, AA237077, AA243142
13	HPTVH24	23	831983	R72195, H96214, W21593, N89670, AA134226, AA134227, AA186685, AA186686, AA586777, AA622860, AA661756, AA740489, AA908358, AI083795, F19158, AI087053, C00663
17	HUSYG26	27	709709	H54036
20	HSKDS47	30	719749	T78923, T78976, H08729, H18949, H40038, H82652, N21057, N23763, N27889, N39362, W78823
21	HSAYR62	31	751035	T90176, R11168, T82011, T82985, T85770, T96656, T96771, T97008, T97123, R13191, R18831, R20605, R43942, R43942, R56736, R56890, H06618, H07127, H19282, H19283, H46323, H59950, H79981, H79982, H88518, H88559, N23329, N25647, N32490, N33290, N34246, N44820, N63908, N77855, N80585, N98279, W05818, W07707, W48686, W78051, W79145, AA026051, AA031370, AA031371, AA043919, AA129595, AA129594, AA188510
23	HSKYR49	33	806548	R10037, R23518, R26882, R27113, R31387, R32884, H17608, R91091, R91092, H79471
25	HUFAU71	35	638836	T95298, T95379, R14499, R22689, R40829, R43395, R43395, R40829, R56598, R56599, R59573, R59632, H16755, H16864, H17318, H17331, H17342, H17360, H29327, H29420, H62754, H99958, N21027, N24761, N27951, N35901, N40717, N53455, N72458, AA027842, AA027907, AA058371, AA058446, AA132582, AA132895, AA132787, AA133062, AA151316, AA151317
26	HUFBK95	36	683039	R35058, AA019290, AA054208, AA056314
28	HUFBV62	38	742892	R55017, AA114978



31	HTXJY08	41	637774	N44737
35	H6EDM64	45	841331	N24911, N41812, N78953, W17347, AA022454, AA022523, AA053148, AA053706, AA112404, AA483525, AA524482, AA622680, AA576098, AA687451, AA687509, AA689495, AA689552, AA745895, AA977756, AA113160, AA480131, AA482098, F22287, AA994188, AI061434
38	HMIAG72	48	824091	R08907, R09013, T77975, T78824, R38406, R64249, H06528, H06583, H16516, H16659, H86591, H87995, H88170, H88170, N22699, N50518, N50574, N53617, N71885, N72008, N99442, W24335, W25168, W81578, W81643, AA121595, AA132604, AA132808, AA164624, AA176369, AA180396, AA181866, AA182552, AA232698, AA232699, AA232927, AA418258, AA418361
39	HISAT67	49	843549	R12741, R13035, R27176, R31157, R37205, R39793, R40419, R40419, R69998, R74188, R74275, H06495, H11618, H58370, H58371, H83961, N35214, N75291, N81134, N99252, N90293, AA058936, AA088263, AA151552, AA151496, AA151497, AA252504, AA252420, AA426271, AA425979, AA548944, AA564447, AA565575, AA580861, AA613326, AA632377, AA687138, AA769320, AA827859, AA877093, AA976201, C02190, AA292350, AA394068, AA411464, AA411591, AA478985, AA477921, AA629936, AA889019, Z17827, AI085046, AI085749, AI089337, AI096361, T25048, AA699494
41	HT5FX79	51	794169	T98950, T99001, H50838, W72985, W76184, AA179812
46	HOUFS04	56	771564	T54671, T54758, R33103, R76568, R76642, H02915, H03822, H28472, H39966, R84685, R84707, R84949, R85648, R85665, H95551, N22080, N52145, N75868, AA011013, AA010836, AA011115, AA016056, AA016111, AA021571, AA020805, AA058678, AA058782, AA419376
47	HOGAH40	57	790978	R00829, R01485, R02125, R02239, R60597, H22159, H22211, H44264, H44593, R86167, H63503, H63903, H99875, N20009, N22676, N25924, N28560, N39271, N70121, N72131, N74513, N78846, W00380, W02880, W04149, W17120, W46894, W61269, W61270, AA100544, AA100543, AA115208, AA115218, AA133166, AA132440, AA132608
50	HMEJK28	60	825591	R17124, R43217, R43217, R59853, R59854, R62245, H82727, H83601, N20100, N23783, N45714, N48992, N51271, N51936, N53390, N67649, N95707, W00832, W01099, W04513, W06962, W31455, W38741, W46970, W47070, W81364, W81623, AA056374, AA056425, AA064891, AA167747, AA167647

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

5

### Examples

#### Example 1: Isolation of a Selected cDNA Clone From the Deposited Sample

Each cDNA clone in a cited ATCC deposit is contained in a plasmid vector. Table 1 identifies the vectors used to construct the cDNA library from which each  
 10 clone was isolated. In many cases, the vector used to construct the library is a phage vector from which a plasmid has been excised. The table immediately below correlates the related plasmid for each phage vector used in constructing the cDNA library. For example, where a particular clone is identified in Table 1 as being isolated in the vector "Lambda Zap," the corresponding deposited clone is in  
 15 "pBluescript."

	<u>Vector Used to Construct Library</u>	<u>Corresponding Deposited</u>
	<u>Plasmid</u>	
	Lambda Zap	pBluescript (pBS)
	Uni-Zap XR	pBluescript (pBS)
20	Zap Express	pBK
	lafmid BA	plafmid BA
	pSport1	pSport1
	pCMVSPORT 2.0	pCMVSPORT 2.0
	pCMVSPORT 3.0	pCMVSPORT 3.0
25	pCR®2.1	pCR®2.1

Vectors Lambda Zap (U.S. Patent Nos. 5,128,256 and 5,286,636), Uni-Zap XR (U.S. Patent Nos. 5,128, 256 and 5,286,636), Zap Express (U.S. Patent Nos. 5,128,256 and 5,286,636), pBluescript (pBS) (Short, J. M. et al., Nucleic Acids Res. 16:7583-7600 (1988); Alting-Mees, M. A. and Short, J. M., Nucleic Acids Res. 17:9494 (1989)) and pBK (Alting-Mees, M. A. et al., Strategies 5:58-61 (1992)) are commercially available from Stratagene Cloning Systems, Inc., 11011 N. Torrey Pines Road, La Jolla, CA, 92037. pBS contains an ampicillin resistance gene and pBK contains a neomycin resistance gene. Both can be transformed into E. coli strain XL-1 Blue, also available from Stratagene. pBS comes in 4 forms SK+, SK-, KS+ and KS. The S and K refers to the orientation of the polylinker to the T7 and T3 primer sequences which flank the polylinker region ("S" is for SacI and "K" is for KpnI which are the first sites on each respective end of the linker). "+" or "-" refer to the orientation of the f1 origin of replication ("ori"), such that in one orientation, single stranded rescue initiated from the f1 ori generates sense strand DNA and in the other, antisense.

Vectors pSport1, pCMVSport 2.0 and pCMVSport 3.0, were obtained from Life Technologies, Inc., P. O. Box 6009, Gaithersburg, MD 20897. All Sport vectors contain an ampicillin resistance gene and may be transformed into E. coli strain DH10B, also available from Life Technologies. (See, for instance, Gruber, C. E., et al., Focus 15:59 (1993).) Vector lafmid BA (Bento Soares, Columbia University, NY) contains an ampicillin resistance gene and can be transformed into E. coli strain XL-1 Blue. Vector pCR<sup>®</sup>2.1, which is available from Invitrogen, 1600 Faraday Avenue, Carlsbad, CA 92008, contains an ampicillin resistance gene and may be transformed into E. coli strain DH10B, available from Life Technologies. (See, for instance, Clark, J. M., Nuc. Acids Res. 16:9677-9686 (1988) and Mead, D. et al., Bio/Technology 9: (1991).) Preferably, a polynucleotide of the present invention

does not comprise the phage vector sequences identified for the particular clone in Table 1, as well as the corresponding plasmid vector sequences designated above.

The deposited material in the sample assigned the ATCC Deposit Number cited in Table 1 for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 50 plasmid DNAs, each containing a different cDNA clone; but such a deposit sample may include plasmids for more or less than 50 cDNA clones, up to about 500 cDNA clones.

Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNAs cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:X.

Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with  $^{32}\text{P}$ - $\gamma$ -ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, NY (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for

bacterial colony screening (e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

Alternatively, two primers of 17-20 nucleotides derived from both ends of the  
5 SEQ ID NO:X (i.e., within the region of SEQ ID NO:X bounded by the 5' NT and  
the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the  
desired cDNA using the deposited cDNA plasmid as a template. The polymerase  
chain reaction is carried out under routine conditions, for instance, in 25 ul of reaction  
mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is  
10 1.5-5 mM MgCl<sub>2</sub>, 0.01% (w/v) gelatin, 20 uM each of dATP, dCTP, dGTP, dTTP, 25  
pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR  
(denaturation at 94 degree C for 1 min; annealing at 55 degree C for 1 min; elongation  
at 72 degree C for 1 min) are performed with a Perkin-Elmer Cetus automated  
thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and  
15 the DNA band with expected molecular weight is excised and purified. The PCR  
product is verified to be the selected sequence by subcloning and sequencing the  
DNA product.

Several methods are available for the identification of the 5' or 3' non-coding  
portions of a gene which may not be present in the deposited clone. These methods  
20 include but are not limited to, filter probing, clone enrichment using specific probes,  
and protocols similar or identical to 5' and 3' "RACE" protocols which are well  
known in the art. For instance, a method similar to 5' RACE is available for  
generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et  
al., *Nucleic Acids Res.* 21(7):1683-1684 (1993).)

25 Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a  
population of RNA presumably containing full-length gene RNA transcripts. A

primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full length gene.

5        This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order  
10    to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

      This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is  
15    used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene.

20    **Example 2: Isolation of Genomic Clones Corresponding to a Polynucleotide**

      A human genomic P1 library (Genomic Systems, Inc.) is screened by PCR using primers selected for the cDNA sequence corresponding to SEQ ID NO:X., according to the method described in Example 1. (See also, Sambrook.)

25    **Example 3: Tissue Distribution of Polypeptide**

Tissue distribution of mRNA expression of polynucleotides of the present invention is determined using protocols for Northern blot analysis, described by, among others, Sambrook et al. For example, a cDNA probe produced by the method described in Example 1 is labeled with  $P^{32}$  using the rediprime™ DNA labeling  
5 system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for mRNA expression.

10 Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) (Clontech) are examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70 degree C overnight, and the films  
15 developed according to standard procedures.

#### **Example 4: Chromosomal Mapping of the Polynucleotides**

An oligonucleotide primer set is designed according to the sequence at the 5' end of SEQ ID NO:X. This primer preferably spans about 100 nucleotides. This  
20 primer set is then used in a polymerase chain reaction under the following set of conditions : 30 seconds, 95 degree C; 1 minute, 56 degree C; 1 minute, 70 degree C. This cycle is repeated 32 times followed by one 5 minute cycle at 70 degree C. Human, mouse, and hamster DNA is used as template in addition to a somatic cell hybrid panel containing individual chromosomes or chromosome fragments (Bios,  
25 Inc). The reactions is analyzed on either 8% polyacrylamide gels or 3.5 % agarose

gels. Chromosome mapping is determined by the presence of an approximately 100 bp PCR fragment in the particular somatic cell hybrid.

**Example 5: Bacterial Expression of a Polypeptide**

5           A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 1, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product  
10 into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, CA). This plasmid vector encodes antibiotic resistance (Amp<sup>r</sup>), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning  
15 sites.

          The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the E. coli strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which  
20 expresses the lacI repressor and also confers kanamycin resistance (Kan<sup>r</sup>). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

          Clones containing the desired constructs are grown overnight (O/N) in liquid  
25 culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The



cells are grown to an optical density 600 (O.D.<sup>600</sup>) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

5           Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000Xg). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4 degree C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from  
10   QIAGEN, Inc., *supra*). Proteins with a 6 x His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., *supra*).

Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then  
15   washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-  
20   NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50  
25   mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4 degree C or frozen at -80 degree C.

In addition to the above expression vector, the present invention further includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains:

- 5 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (*lacIq*). The origin of replication (*oriC*) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences are made synthetically.

- 10 DNA can be inserted into the pHEa by restricting the vector with *NdeI* and *XbaI*, *BamHI*, *XhoI*, or *Asp718*, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for *NdeI* (5' primer) and *XbaI*, *BamHI*, *XhoI*, or  
15 *Asp718* (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols.

The engineered vector could easily be substituted in the above protocol to express protein in a bacterial system.

20 **Example 6: Purification of a Polypeptide from an Inclusion Body**

The following alternative method can be used to purify a polypeptide expressed in *E. coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10 degree C.

- 25 Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10 degree C and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield

of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

5           The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

10           The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000 xg centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4 degree C overnight to allow further GuHCl extraction.

            Following high speed centrifugation (30,000 xg) to remove insoluble particles,  
15           the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4 degree C without mixing for 12 hours prior to further purification steps.

            To clarify the refolded polypeptide solution, a previously prepared tangential  
20           filtration unit equipped with 0.16 um membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a  
25           stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are  
5 equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant  $A_{280}$  monitoring of the effluent. Fractions containing the  
10 polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 ug of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS  
15 contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

### **Example 7: Cloning and Expression of a Polypeptide in a Baculovirus**

#### **Expression System**

20 In this example, the plasmid shuttle vector pA2 is used to insert a polynucleotide into a baculovirus to express a polypeptide. This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40")  
25 is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak

Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

5           Many other baculovirus vectors can be used in place of the vector above, such as pAc373, pVL941, and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-  
10   39 (1989).

          Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified in Table 1, is amplified using the PCR protocol described in Example 1. If the naturally occurring signal sequence is used to produce the secreted protein, the pA2 vector does  
15   not need a second signal peptide. Alternatively, the vector can be modified (pA2 GP) to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

20           The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

          The plasmid is digested with the corresponding restriction enzymes and  
25   optionally, can be dephosphorylated using calf intestinal phosphatase, using routine

procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("GeneClean" BIO 101 Inc., La Jolla, Ca.).

The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

Five ug of a plasmid containing the polynucleotide is co-transfected with 1.0 ug of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One ug of BaculoGold™ virus DNA and 5 ug of the plasmid are mixed in a sterile well of a microtiter plate containing 50 ul of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 ul Lipofectin plus 90 ul Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27 degrees C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27 degrees C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of

a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then  
5 resuspended in a microcentrifuge tube containing 200 ul of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4 degree C.

To verify the expression of the polypeptide, Sf9 cells are grown in Grace's  
10 medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5  
15 uCi of  $^{35}\text{S}$ -methionine and 5 uCi  $^{35}\text{S}$ -cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

Microsequencing of the amino acid sequence of the amino terminus of  
20 purified protein may be used to determine the amino terminal sequence of the produced protein.

#### **Example 8: Expression of a Polypeptide in Mammalian Cells**

The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which  
25 mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the

transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSPORT 2.0, and pCMVSPORT 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest



resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

Derivatives of the plasmid pSV2-dhfr (ATCC Accession No. 37146), the  
5 expression vectors pC4 (ATCC Accession No. 209646) and pC6 (ATCC Accession  
No.209647) contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et  
al., Molecular and Cellular Biology, 438-447 (March, 1985)) plus a fragment of the  
CMV-enhancer (Boshart et al., Cell 41:521-530 (1985).) Multiple cloning sites, e.g.,  
with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the  
10 cloning of the gene of interest. The vectors also contain the 3' intron, the  
polyadenylation and termination signal of the rat preproinsulin gene, and the mouse  
DHFR gene under control of the SV40 early promoter.

Specifically, the plasmid pC6, for example, is digested with appropriate  
restriction enzymes and then dephosphorylated using calf intestinal phosphates by  
15 procedures known in the art. The vector is then isolated from a 1% agarose gel.

A polynucleotide of the present invention is amplified according to the  
protocol outlined in Example 1. If the naturally occurring signal sequence is used to  
produce the secreted protein, the vector does not need a second signal peptide.  
Alternatively, if the naturally occurring signal sequence is not used, the vector can be  
20 modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

The amplified fragment is isolated from a 1% agarose gel using a  
commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment  
then is digested with appropriate restriction enzymes and again purified on a 1%  
agarose gel.

25 The amplified fragment is then digested with the same restriction enzyme and  
purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector

are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

Chinese hamster ovary cells lacking an active DHFR gene is used for  
5 transfection. Five  $\mu$ g of the expression plasmid pC6 a pC4 is cotransfected with 0.5  
ug of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-  
neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an  
enzyme that confers resistance to a group of antibiotics including G418. The cells are  
seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the  
10 cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in  
alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1  
mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in  
6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50  
nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest  
15 concentrations of methotrexate are then transferred to new 6-well plates containing  
even higher concentrations of methotrexate (1 uM, 2 uM, 5 uM, 10 mM, 20 mM).  
The same procedure is repeated until clones are obtained which grow at a  
concentration of 100 - 200 uM. Expression of the desired gene product is analyzed,  
for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

20

#### **Example 9: Protein Fusions**

The polypeptides of the present invention are preferably fused to other  
proteins. These fusion proteins can be used for a variety of applications. For  
example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG  
25 domains, and maltose binding protein facilitates purification. (See Example 5; see  
also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to

IgG-1, IgG-3, and albumin increases the halflife time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create  
5 chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule, or the protocol described in Example 5.

10 Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector.

For example, if pC4 (Accession No. 209646) is used, the human Fc portion  
15 can be ligated into the BamHI cloning site. Note that the 3' BamHI site should be destroyed. Next, the vector containing the human Fc portion is re-restricted with BamHI, linearizing the vector, and a polynucleotide of the present invention, isolated by the PCR protocol described in Example 1, is ligated into this BamHI site. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will  
20 not be produced.

If the naturally occurring signal sequence is used to produce the secreted protein, pC4 does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.)

25

Human IgG Fc region:

GGGATCCGGAGCCCAAATCTTCTGACAAAACCTCACACATGCCCACCGTGC  
CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCAGAAA  
CCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTACATGCGTGGT  
GGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGG  
5 ACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA  
CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT  
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA  
ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC  
CACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAG  
10 GTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGT  
GGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCT  
CCCGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTG  
GACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCA  
TGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG  
15 GTAAATGAGTGCGACGGCCGCGACTCTAGAGGAT (SEQ ID NO:1)

**Example 10: Production of an Antibody from a Polypeptide**

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) As one example of such methods, cells  
20 expressing a polypeptide of the present invention is administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of the secreted protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

25 In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or protein binding fragments thereof). Such monoclonal

antibodies can be prepared using hybridoma technology. (Köhler et al., Nature 256:495 (1975); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981).) In general, such procedures  
5 involve immunizing an animal (preferably a mouse) with polypeptide or, more preferably, with a secreted polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56 degrees C), and supplemented with about 10 g/l of nonessential amino acids,  
10 about 1,000 U/ml of penicillin, and about 100 ug/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are  
15 selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands et al. (Gastroenterology 80:225-232 (1981).) The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide.

Alternatively, additional antibodies capable of binding to the polypeptide can  
20 be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells,  
25 and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the

polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

It will be appreciated that Fab and F(ab')<sub>2</sub> and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')<sub>2</sub> fragments). Alternatively, secreted protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

For in vivo use of antibodies in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

20

#### **Example 11: Production Of Secreted Protein For High-Throughput Screening Assays**

The following protocol produces a supernatant containing a polypeptide to be tested. This supernatant can then be used in the Screening Assays described in Examples 13-20.

25

First, dilute Poly-D-Lysine (644 587 Boehringer-Mannheim) stock solution (1mg/ml in PBS) 1:20 in PBS (w/o calcium or magnesium 17-516F Biowhittaker) for a working solution of 50ug/ml. Add 200 ul of this solution to each well (24 well plates) and incubate at RT for 20 minutes. Be sure to distribute the solution over each well (note: a 12-channel pipetter may be used with tips on every other channel). Aspirate off the Poly-D-Lysine solution and rinse with 1ml PBS (Phosphate Buffered Saline). The PBS should remain in the well until just prior to plating the cells and plates may be poly-lysine coated in advance for up to two weeks.

Plate 293T cells (do not carry cells past P+20) at  $2 \times 10^5$  cells/well in .5ml DMEM(Dulbecco's Modified Eagle Medium)(with 4.5 G/L glucose and L-glutamine (12-604F Biowhittaker))/10% heat inactivated FBS(14-503F Biowhittaker)/1x Penstrep(17-602E Biowhittaker). Let the cells grow overnight.

The next day, mix together in a sterile solution basin: 300 ul Lipofectamine (18324-012 Gibco/BRL) and 5ml Optimem I (31985070 Gibco/BRL)/96-well plate. With a small volume multi-channel pipetter, aliquot approximately 2ug of an expression vector containing a polynucleotide insert, produced by the methods described in Examples 8 or 9, into an appropriately labeled 96-well round bottom plate. With a multi-channel pipetter, add 50ul of the Lipofectamine/Optimem I mixture to each well. Pipette up and down gently to mix. Incubate at RT 15-45 minutes. After about 20 minutes, use a multi-channel pipetter to add 150ul Optimem I to each well. As a control, one plate of vector DNA lacking an insert should be transfected with each set of transfections.

Preferably, the transfection should be performed by tag-teaming the following tasks. By tag-teaming, hands on time is cut in half, and the cells do not spend too much time on PBS. First, person A aspirates off the media from four 24-well plates of cells, and then person B rinses each well with .5-1ml PBS. Person A then aspirates

off PBS rinse, and person B, using a 12-channel pipetter with tips on every other channel, adds the 200ul of DNA/Lipofectamine/Optimem I complex to the odd wells first, then to the even wells, to each row on the 24-well plates. Incubate at 37 degrees C for 6 hours.

- 5           While cells are incubating, prepare appropriate media, either 1%BSA in DMEM with 1x penstrep, or CHO-5 media (116.6 mg/L of  $\text{CaCl}_2$  (anhyd); 0.00130 mg/L  $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ; 0.050 mg/L of  $\text{Fe}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$ ; 0.417 mg/L of  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ; 311.80 mg/L of Kcl; 28.64 mg/L of  $\text{MgCl}_2$ ; 48.84 mg/L of  $\text{MgSO}_4$ ; 6995.50 mg/L of NaCl; 2400.0 mg/L of  $\text{NaHCO}_3$ ; 62.50 mg/L of  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ ; 71.02 mg/L of
- 10  $\text{Na}_2\text{HPO}_4$ ; .4320 mg/L of  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ ; .002 mg/L of Arachidonic Acid ; 1.022 mg/L of Cholesterol; .070 mg/L of DL-alpha-Tocopherol-Acetate; 0.0520 mg/L of Linoleic Acid; 0.010 mg/L of Linolenic Acid; 0.010 mg/L of Myristic Acid; 0.010 mg/L of Oleic Acid; 0.010 mg/L of Palmitric Acid; 0.010 mg/L of Palmitic Acid; 100 mg/L of Pluronic F-68; 0.010 mg/L of Stearic Acid; 2.20 mg/L of Tween 80; 4551 mg/L of D-
- 15 Glucose; 130.85 mg/ml of L- Alanine; 147.50 mg/ml of L-Arginine-HCL; 7.50 mg/ml of L-Asparagine- $\text{H}_2\text{O}$ ; 6.65 mg/ml of L-Aspartic Acid; 29.56 mg/ml of L-Cystine-2HCL- $\text{H}_2\text{O}$ ; 31.29 mg/ml of L-Cystine-2HCL; 7.35 mg/ml of L-Glutamic Acid; 365.0 mg/ml of L-Glutamine; 18.75 mg/ml of Glycine; 52.48 mg/ml of L-Histidine-HCL- $\text{H}_2\text{O}$ ; 106.97 mg/ml of L-Isoleucine; 111.45 mg/ml of L-Leucine; 163.75 mg/ml of L-
- 20 Lysine HCL; 32.34 mg/ml of L-Methionine; 68.48 mg/ml of L-Phenylalainine; 40.0 mg/ml of L-Proline; 26.25 mg/ml of L-Serine; 101.05 mg/ml of L-Threonine; 19.22 mg/ml of L-Tryptophan; 91.79 mg/ml of L-Tyrosine-2Na-2 $\text{H}_2\text{O}$ ; 99.65 mg/ml of L-Valine; 0.0035 mg/L of Biotin; 3.24 mg/L of D-Ca Pantothenate; 11.78 mg/L of Choline Chloride; 4.65 mg/L of Folic Acid; 15.60 mg/L of i-Inositol; 3.02 mg/L of
- 25 Niacinamide; 3.00 mg/L of Pyridoxal HCL; 0.031 mg/L of Pyridoxine HCL; 0.319 mg/L of Riboflavin; 3.17 mg/L of Thiamine HCL; 0.365 mg/L of Thymidine; and



0.680 mg/L of Vitamin B<sub>12</sub>; 25 mM of HEPES Buffer; 2.39 mg/L of Na Hypoxanthine; 0.105 mg/L of Lipoic Acid; 0.081 mg/L of Sodium Putrescine-2HCL; 55.0 mg/L of Sodium Pyruvate; 0.0067 mg/L of Sodium Selenite; 20uM of Ethanolamine; 0.122 mg/L of Ferric Citrate; 41.70 mg/L of Methyl-B-Cyclodextrin  
5 complexed with Linoleic Acid; 33.33 mg/L of Methyl-B-Cyclodextrin complexed with Oleic Acid; and 10 mg/L of Methyl-B-Cyclodextrin complexed with Retinal) with 2mm glutamine and 1x penstrep. (BSA (81-068-3 Bayer) 100gm dissolved in 1L DMEM for a 10% BSA stock solution). Filter the media and collect 50 ul for endotoxin assay in 15ml polystyrene conical.

10 The transfection reaction is terminated, preferably by tag-teaming, at the end of the incubation period. Person A aspirates off the transfection media, while person B adds 1.5ml appropriate media to each well. Incubate at 37 degrees C for 45 or 72 hours depending on the media used: 1%BSA for 45 hours or CHO-5 for 72 hours.

On day four, using a 300ul multichannel pipetter, aliquot 600ul in one 1ml  
15 deep well plate and the remaining supernatant into a 2ml deep well. The supernatants from each well can then be used in the assays described in Examples 13-20.

It is specifically understood that when activity is obtained in any of the assays described below using a supernatant, the activity originates from either the polypeptide directly (e.g., as a secreted protein) or by the polypeptide inducing  
20 expression of other proteins, which are then secreted into the supernatant. Thus, the invention further provides a method of identifying the protein in the supernatant characterized by an activity in a particular assay.

#### **Example 12: Construction of GAS Reporter Construct**

25 One signal transduction pathway involved in the differentiation and proliferation of cells is called the Jaks-STATs pathway. Activated proteins in the

Jaks-STATs pathway bind to gamma activation site “GAS” elements or interferon-sensitive responsive element (“ISRE”), located in the promoter of many genes. The binding of a protein to these elements alter the expression of the associated gene.

GAS and ISRE elements are recognized by a class of transcription factors called Signal Transducers and Activators of Transcription, or “STATs.” There are six members of the STATs family. Stat1 and Stat3 are present in many cell types, as is Stat2 (as response to IFN-alpha is widespread). Stat4 is more restricted and is not in many cell types though it has been found in T helper class I, cells after treatment with IL-12. Stat5 was originally called mammary growth factor, but has been found at higher concentrations in other cells including myeloid cells. It can be activated in tissue culture cells by many cytokines.

The STATs are activated to translocate from the cytoplasm to the nucleus upon tyrosine phosphorylation by a set of kinases known as the Janus Kinase (“Jaks”) family. Jaks represent a distinct family of soluble tyrosine kinases and include Tyk2, Jak1, Jak2, and Jak3. These kinases display significant sequence similarity and are generally catalytically inactive in resting cells.

The Jaks are activated by a wide range of receptors summarized in the Table below. (Adapted from review by Schidler and Darnell, Ann. Rev. Biochem. 64:621-51 (1995).) A cytokine receptor family, capable of activating Jaks, is divided into two groups: (a) Class 1 includes receptors for IL-2, IL-3, IL-4, IL-6, IL-7, IL-9, IL-11, IL-12, IL-15, Epo, PRL, GH, G-CSF, GM-CSF, LIF, CNTF, and thrombopoietin; and (b) Class 2 includes IFN-a, IFN-g, and IL-10. The Class 1 receptors share a conserved cysteine motif (a set of four conserved cysteines and one tryptophan) and a WSXWS motif (a membrane proximal region encoding Trp-Ser-Xxx-Trp-Ser (SEQ ID NO:2)).

Thus, on binding of a ligand to a receptor, Jaks are activated, which in turn activate STATs, which then translocate and bind to GAS elements. This entire process is encompassed in the Jaks-STATs signal transduction pathway.

Therefore, activation of the Jaks-STATs pathway, reflected by the binding of  
5 the GAS or the ISRE element, can be used to indicate proteins involved in the proliferation and differentiation of cells. For example, growth factors and cytokines are known to activate the Jaks-STATs pathway. (See Table below.) Thus, by using GAS elements linked to reporter molecules, activators of the Jaks-STATs pathway can be identified.

<u>Ligand</u>	<u>JAKs</u>				<u>STATS</u>	<u>GAS(elements) or ISRE</u>
	<u>tyk2</u>	<u>Jak1</u>	<u>Jak2</u>	<u>Jak3</u>		
<u>IFN family</u>						
IFN-a/B	+	+	-	-	1,2,3	ISRE
IFN-g		+	+	-	1	GAS (IRF1>Lys6>IFP)
Il-10	+	?	?	-	1,3	
<u>gp130 family</u>						
IL-6 (Pleiotrophic)	+	+	+	?	1,3	GAS (IRF1>Lys6>IFP)
Il-11(Pleiotrophic)	?	+	?	?	1,3	
OnM(Pleiotrophic)	?	+	+	?	1,3	
LIF(Pleiotrophic)	?	+	+	?	1,3	
CNTF(Pleiotrophic)	-/+	+	+	?	1,3	
G-CSF(Pleiotrophic)	?	+	?	?	1,3	
IL-12(Pleiotrophic)	+	-	+	+	1,3	
<u>g-C family</u>						
IL-2 (lymphocytes)	-	+	-	+	1,3,5	GAS
IL-4 (lymph/myeloid)	-	+	-	+	6	GAS (IRF1 = IFP >>Ly6)(IgH)
IL-7 (lymphocytes)	-	+	-	+	5	GAS
IL-9 (lymphocytes)	-	+	-	+	5	GAS
IL-13 (lymphocyte)	-	+	?	?	6	GAS
IL-15	?	+	?	+	5	GAS
<u>gp140 family</u>						
IL-3 (myeloid)	-	-	+	-	5	GAS (IRF1>IFP>>Ly6)
IL-5 (myeloid)	-	-	+	-	5	GAS
GM-CSF (myeloid)	-	-	+	-	5	GAS
<u>Growth hormone family</u>						
GH	?	-	+	-	5	
PRL	?	+/-	+	-	1,3,5	
EPO	?	-	+	-	5	GAS(B-CAS>IRF1=IFP>>Ly6)
<u>Receptor Tyrosine Kinases</u>						
EGF	?	+	+	-	1,3	GAS (IRF1)
PDGF	?	+	+	-	1,3	
CSF-1	?	+	+	-	1,3	GAS (not IRF1)

To construct a synthetic GAS containing promoter element, which is used in the Biological Assays described in Examples 13-14, a PCR based strategy is employed to generate a GAS-SV40 promoter sequence. The 5' primer contains four tandem copies of the GAS binding site found in the JRF1 promoter and previously demonstrated to bind STATs upon induction with a range of cytokines (Rothman et al., Immunity 1:457-468 (1994).), although other GAS or ISRE elements can be used instead. The 5' primer also contains 18bp of sequence complementary to the SV40 early promoter sequence and is flanked with an XhoI site. The sequence of the 5' primer is:

10 5':GCGCCTCGAGATTTCCTCCCGAAATCTAGATTTCCTCCCGAAATGATTTCCTCCCGAAATGATTTCCTCCCGAAATATCTGCCATCTCAATTAG:3' (SEQ ID NO:3)

The downstream primer is complementary to the SV40 promoter and is flanked with a Hind III site: 5':GCGGCAAGCTTTTGTCAAAGCCTAGGC:3' (SEQ ID NO:4)

15 PCR amplification is performed using the SV40 promoter template present in the B-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI/Hind III and subcloned into BLSK2-. (Stratagene.) Sequencing with forward and reverse primers confirms that the insert contains the following sequence:

20 5':CTCGAGATTTCCTCCCGAAATCTAGATTTCCTCCCGAAATGATTTCCTCCCGAAATGATTTCCTCCCGAAATATCTGCCATCTCAATTAGTCAGCAACCATAGTCCCGCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCC  
TCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCT  
25 AGGCTTTTGCAAAAAGCTT:3' (SEQ ID NO:5)

With this GAS promoter element linked to the SV40 promoter, a GAS:SEAP2 reporter construct is next engineered. Here, the reporter molecule is a secreted alkaline phosphatase, or "SEAP." Clearly, however, any reporter molecule can be instead of SEAP, in this or in any of the other Examples. Well known reporter  
5 molecules that can be used instead of SEAP include chloramphenicol acetyltransferase (CAT), luciferase, alkaline phosphatase, B-galactosidase, green fluorescent protein (GFP), or any protein detectable by an antibody.

The above sequence confirmed synthetic GAS-SV40 promoter element is subcloned into the pSEAP-Promoter vector obtained from Clontech using HindIII and  
10 XhoI, effectively replacing the SV40 promoter with the amplified GAS:SV40 promoter element, to create the GAS-SEAP vector. However, this vector does not contain a neomycin resistance gene, and therefore, is not preferred for mammalian expression systems.

Thus, in order to generate mammalian stable cell lines expressing the GAS-  
15 SEAP reporter, the GAS-SEAP cassette is removed from the GAS-SEAP vector using SalI and NotI, and inserted into a backbone vector containing the neomycin resistance gene, such as pGFP-1 (Clontech), using these restriction sites in the multiple cloning site, to create the GAS-SEAP/Neo vector. Once this vector is transfected into mammalian cells, this vector can then be used as a reporter molecule for GAS binding  
20 as described in Examples 13-14.

Other constructs can be made using the above description and replacing GAS with a different promoter sequence. For example, construction of reporter molecules containing NFK-B and EGR promoter sequences are described in Examples 15 and  
16. However, many other promoters can be substituted using the protocols described  
25 in these Examples. For instance, SRE, IL-2, NFAT, or Osteocalcin promoters can be substituted, alone or in combination (e.g., GAS/NF-KB/EGR, GAS/NF-KB, IL-

2/NFAT, or NF-KB/GAS). Similarly, other cell lines can be used to test reporter construct activity, such as HELA (epithelial), HUVEC (endothelial), Reh (B-cell), Saos-2 (osteoblast), HUVAC (aortic), or Cardiomyocyte.

5 **Example 13: High-Throughput Screening Assay for T-cell Activity.**

The following protocol is used to assess T-cell activity by identifying factors, and determining whether supernate containing a polypeptide of the invention proliferates and/or differentiates T-cells. T-cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP  
10 activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The T-cell used in this assay is Jurkat T-cells (ATCC Accession No. TIB-152), although Molt-3 cells (ATCC Accession No. CRL-1552) and Molt-4 cells (ATCC Accession No. CRL-1582) cells can also be used.

Jurkat T-cells are lymphoblastic CD4+ Th1 helper cells. In order to generate  
15 stable cell lines, approximately 2 million Jurkat cells are transfected with the GAS-SEAP/neo vector using DMRIE-C (Life Technologies)(transfection procedure described below). The transfected cells are seeded to a density of approximately 20,000 cells per well and transfectants resistant to 1 mg/ml gentamicin selected. Resistant colonies are expanded and then tested for their response to increasing  
20 concentrations of interferon gamma. The dose response of a selected clone is demonstrated.

Specifically, the following protocol will yield sufficient cells for 75 wells containing 200 ul of cells. Thus, it is either scaled up, or performed in multiple to generate sufficient cells for multiple 96 well plates. Jurkat cells are maintained in  
25 RPMI + 10% serum with 1%Pen-Strep. Combine 2.5 mls of OPTI-MEM (Life

Technologies) with 10 ug of plasmid DNA in a T25 flask. Add 2.5 ml OPTI-MEM containing 50 ul of DMRIE-C and incubate at room temperature for 15-45 mins.

During the incubation period, count cell concentration, spin down the required number of cells ( $10^7$  per transfection), and resuspend in OPTI-MEM to a final  
5 concentration of  $10^7$  cells/ml. Then add 1ml of  $1 \times 10^7$  cells in OPTI-MEM to T25 flask and incubate at 37 degrees C for 6 hrs. After the incubation, add 10 ml of RPMI + 15% serum.

The Jurkat:GAS-SEAP stable reporter lines are maintained in RPMI + 10% serum, 1 mg/ml Genticin, and 1% Pen-Strep. These cells are treated with  
10 supernatants containing polypeptides of the invention and/or induced polypeptides of the invention as produced by the protocol described in Example 11.

On the day of treatment with the supernatant, the cells should be washed and resuspended in fresh RPMI + 10% serum to a density of 500,000 cells per ml. The exact number of cells required will depend on the number of supernatants being  
15 screened. For one 96 well plate, approximately 10 million cells (for 10 plates, 100 million cells) are required.

Transfer the cells to a triangular reservoir boat, in order to dispense the cells into a 96 well dish, using a 12 channel pipette. Using a 12 channel pipette, transfer 200 ul of cells into each well (therefore adding 100, 000 cells per well).

20 After all the plates have been seeded, 50 ul of the supernatants are transferred directly from the 96 well plate containing the supernatants into each well using a 12 channel pipette. In addition, a dose of exogenous interferon gamma (0.1, 1.0, 10 ng) is added to wells H9, H10, and H11 to serve as additional positive controls for the assay.

25 The 96 well dishes containing Jurkat cells treated with supernatants are placed in an incubator for 48 hrs (note: this time is variable between 48-72 hrs). 35 ul



samples from each well are then transferred to an opaque 96 well plate using a 12 channel pipette. The opaque plates should be covered (using sellophane covers) and stored at -20 degrees C until SEAP assays are performed according to Example 17. The plates containing the remaining treated cells are placed at 4 degrees C and serve  
5 as a source of material for repeating the assay on a specific well if desired.

As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate Jurkat T cells. Over 30 fold induction is typically observed in the positive control wells.

The above protocol may be used in the generation of both transient, as well as,  
10 stable transfected cells, which would be apparent to those of skill in the art.

#### **Example 14: High-Throughput Screening Assay Identifying Myeloid Activity**

The following protocol is used to assess myeloid activity by determining whether polypeptides of the invention proliferates and/or differentiates myeloid cells.  
15 Myeloid cell activity is assessed using the GAS/SEAP/Neo construct produced in Example 12. Thus, factors that increase SEAP activity indicate the ability to activate the Jaks-STATS signal transduction pathway. The myeloid cell used in this assay is U937, a pre-monocyte cell line, although TF-1, HL60, or KG1 can be used.

To transiently transfect U937 cells with the GAS/SEAP/Neo construct  
20 produced in Example 12, a DEAE-Dextran method (Kharbanda et. al., 1994, Cell Growth & Differentiation, 5:259-265) is used. First, harvest  $2 \times 10^7$  U937 cells and wash with PBS. The U937 cells are usually grown in RPMI 1640 medium containing 10% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 mg/ml streptomycin.

25 Next, suspend the cells in 1 ml of 20 mM Tris-HCl (pH 7.4) buffer containing 0.5 mg/ml DEAE-Dextran, 8 ug GAS-SEAP2 plasmid DNA, 140 mM NaCl, 5 mM

KCl, 375  $\mu$ M  $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 1 mM  $\text{MgCl}_2$ , and 675  $\mu$ M  $\text{CaCl}_2$ . Incubate at 37 degrees C for 45 min.

Wash the cells with RPMI 1640 medium containing 10% FBS and then resuspend in 10 ml complete medium and incubate at 37 degrees C for 36 hr.

- 5        The GAS-SEAP/U937 stable cells are obtained by growing the cells in 400  $\mu$ g/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 400  $\mu$ g/ml G418 for couple of passages.

- 10       These cells are tested by harvesting  $1 \times 10^8$  cells (this is enough for ten 96-well plates assay) and wash with PBS. Suspend the cells in 200 ml above described growth medium, with a final density of  $5 \times 10^5$  cells/ml. Plate 200  $\mu$ l cells per well in the 96-well plate (or  $1 \times 10^5$  cells/well).

- 15       Add 50  $\mu$ l of the supernatant prepared by the protocol described in Example 11. Incubate at 37 degrees C for 48 to 72 hr. As a positive control, 100 Unit/ml interferon gamma can be used which is known to activate U937 cells. Over 30 fold induction is typically observed in the positive control wells. SEAP assay the supernatant according to the protocol described in Example 17.

**Example 15: High-Throughput Screening Assay Identifying Neuronal Activity.**

- 20       When cells undergo differentiation and proliferation, a group of genes are activated through many different signal transduction pathways. One of these genes, EGR1 (early growth response gene 1), is induced in various tissues and cell types upon activation. The promoter of EGR1 is responsible for such induction. Using the EGR1 promoter linked to reporter molecules, activation of cells can be assessed.

- 25       Particularly, the following protocol is used to assess neuronal activity in PC12 cell lines. PC12 cells (rat pheochromocytoma cells) are known to proliferate and/or differentiate by activation with a number of mitogens, such as TPA (tetradecanoyl

phorbol acetate), NGF (nerve growth factor), and EGF (epidermal growth factor).

The EGR1 gene expression is activated during this treatment. Thus, by stably transfecting PC12 cells with a construct containing an EGR promoter linked to SEAP reporter, activation of PC12 cells can be assessed.

5           The EGR/SEAP reporter construct can be assembled by the following protocol. The EGR-1 promoter sequence (-633 to +1)(Sakamoto K et al., Oncogene 6:867-871 (1991)) can be PCR amplified from human genomic DNA using the following primers:

5' GCGCTCGAGGGATGACAGCGATAGAACCCCGG -3' (SEQ ID NO:6)

10           5' GCGAAGCTTCGCGACTCCCCGGATCCGCCTC-3' (SEQ ID NO:7)

Using the GAS:SEAP/Neo vector produced in Example 12, EGR1 amplified product can then be inserted into this vector. Linearize the GAS:SEAP/Neo vector using restriction enzymes XhoI/HindIII, removing the GAS/SV40 stuffer. Restrict the EGR1 amplified product with these same enzymes. Ligate the vector and the EGR1  
15 promoter.

To prepare 96 well-plates for cell culture, two mls of a coating solution (1:30 dilution of collagen type I (Upstate Biotech Inc. Cat#08-115) in 30% ethanol (filter sterilized)) is added per one 10 cm plate or 50 ml per well of the 96-well plate, and allowed to air dry for 2 hr.

20           PC12 cells are routinely grown in RPMI-1640 medium (Bio Whittaker) containing 10% horse serum (JRH BIOSCIENCES, Cat. # 12449-78P), 5% heat-inactivated fetal bovine serum (FBS) supplemented with 100 units/ml penicillin and 100 ug/ml streptomycin on a precoated 10 cm tissue culture dish. One to four split is done every three to four days. Cells are removed from the plates by scraping and  
25 resuspended with pipetting up and down for more than 15 times.

Transfect the EGR/SEAP/Neo construct into PC12 using the Lipofectamine protocol described in Example 11. EGR-SEAP/PC12 stable cells are obtained by growing the cells in 300 ug/ml G418. The G418-free medium is used for routine growth but every one to two months, the cells should be re-grown in 300 ug/ml G418 for couple of passages.

To assay for neuronal activity, a 10 cm plate with cells around 70 to 80% confluent is screened by removing the old medium. Wash the cells once with PBS (Phosphate buffered saline). Then starve the cells in low serum medium (RPMI-1640 containing 1% horse serum and 0.5% FBS with antibiotics) overnight.

The next morning, remove the medium and wash the cells with PBS. Scrape off the cells from the plate, suspend the cells well in 2 ml low serum medium. Count the cell number and add more low serum medium to reach final cell density as  $5 \times 10^5$  cells/ml.

Add 200 ul of the cell suspension to each well of 96-well plate (equivalent to  $1 \times 10^5$  cells/well). Add 50 ul supernatant produced by Example 11, 37°C for 48 to 72 hr. As a positive control, a growth factor known to activate PC12 cells through EGR can be used, such as 50 ng/ul of Neuronal Growth Factor (NGF). Over fifty-fold induction of SEAP is typically seen in the positive control wells. SEAP assay the supernatant according to Example 17.

#### **Example 16: High-Throughput Screening Assay for T-cell Activity**

NF-KB (Nuclear Factor KB) is a transcription factor activated by a wide variety of agents including the inflammatory cytokines IL-1 and TNF, CD30 and CD40, lymphotoxin-alpha and lymphotoxin-beta, by exposure to LPS or thrombin, and by expression of certain viral gene products. As a transcription factor, NF-KB regulates the expression of genes involved in immune cell activation, control of

apoptosis (NF- KB appears to shield cells from apoptosis), B and T-cell development, anti-viral and antimicrobial responses, and multiple stress responses.

In non-stimulated conditions, NF- KB is retained in the cytoplasm with I-KB (Inhibitor KB). However, upon stimulation, I- KB is phosphorylated and degraded,  
5 causing NF- KB to shuttle to the nucleus, thereby activating transcription of target genes. Target genes activated by NF- KB include IL-2, IL-6, GM-CSF, ICAM-1 and class I MHC.

Due to its central role and ability to respond to a range of stimuli, reporter constructs utilizing the NF-KB promoter element are used to screen the supernatants  
10 produced in Example 11. Activators or inhibitors of NF-KB would be useful in treating diseases. For example, inhibitors of NF-KB could be used to treat those diseases related to the acute or chronic activation of NF-KB, such as rheumatoid arthritis.

To construct a vector containing the NF-KB promoter element, a PCR based  
15 strategy is employed. The upstream primer contains four tandem copies of the NF-KB binding site (GGGGACTTTCCC) (SEQ ID NO:8), 18 bp of sequence complementary to the 5' end of the SV40 early promoter sequence, and is flanked with an XhoI site:

5':GCGGCCTCGAGGGGACTTTCCCGGGGACTTTCCGGGGACTTTCCGGGAC  
20 TTTCCATCCTGCCATCTCAATTAG:3' (SEQ ID NO:9)

The downstream primer is complementary to the 3' end of the SV40 promoter and is flanked with a Hind III site:

5':GCGGCAAGCTTTTTGCAAAGCCTAGGC:3' (SEQ ID NO:4)

PCR amplification is performed using the SV40 promoter template present in  
25 the pB-gal:promoter plasmid obtained from Clontech. The resulting PCR fragment is digested with XhoI and Hind III and subcloned into BLSK2-. (Stratagene)

Sequencing with the T7 and T3 primers confirms the insert contains the following sequence:

5' :CTCGAGGGGACTTCCCGGGGACTTCCCGGGGACTTCCCGGGGACTTCCATCTGCCATCTCAATTA  
 5 GTCAGCAACCATAGTCCCGCCCCCTAACTCCGCCCATCCCGCCCCCTAACTCCGCCCAGTTCGCCCATTC  
 TCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATT  
 CCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTT:3' (SEQ ID  
 NO:10)

10           Next, replace the SV40 minimal promoter element present in the pSEAP2-  
 promoter plasmid (Clontech) with this NF-KB/SV40 fragment using XhoI and  
 HindIII. However, this vector does not contain a neomycin resistance gene, and  
 therefore, is not preferred for mammalian expression systems.

          In order to generate stable mammalian cell lines, the NF-KB/SV40/SEAP  
 15 cassette is removed from the above NF-KB/SEAP vector using restriction enzymes  
 Sall and NotI, and inserted into a vector containing neomycin resistance. Particularly,  
 the NF-KB/SV40/SEAP cassette was inserted into pGFP-1 (Clontech), replacing the  
 GFP gene, after restricting pGFP-1 with Sall and NotI.

          Once NF-KB/SV40/SEAP/Neo vector is created, stable Jurkat T-cells are  
 20 created and maintained according to the protocol described in Example 13. Similarly,  
 the method for assaying supernatants with these stable Jurkat T-cells is also described  
 in Example 13. As a positive control, exogenous TNF alpha (0.1, 1, 10 ng) is added to  
 wells H9, H10, and H11, with a 5-10 fold activation typically observed.

#### 25   **Example 17: Assay for SEAP Activity**

          As a reporter molecule for the assays described in Examples 13-16, SEAP  
 activity is assayed using the Tropix Phospho-light Kit (Cat. BP-400) according to the

following general procedure. The Tropix Phospho-light Kit supplies the Dilution, Assay, and Reaction Buffers used below.

Prime a dispenser with the 2.5x Dilution Buffer and dispense 15 ul of 2.5x dilution buffer into Optiplates containing 35 ul of a supernatant. Seal the plates with a plastic sealer and incubate at 65 degree C for 30 min. Separate the Optiplates to avoid uneven heating.

Cool the samples to room temperature for 15 minutes. Empty the dispenser and prime with the Assay Buffer. Add 50 ml Assay Buffer and incubate at room temperature 5 min. Empty the dispenser and prime with the Reaction Buffer (see the table below). Add 50 ul Reaction Buffer and incubate at room temperature for 20 minutes. Since the intensity of the chemiluminescent signal is time dependent, and it takes about 10 minutes to read 5 plates on luminometer, one should treat 5 plates at each time and start the second set 10 minutes later.

Read the relative light unit in the luminometer. Set H12 as blank, and print the results. An increase in chemiluminescence indicates reporter activity.

**Reaction Buffer Formulation:**

# of plates	Rxn buffer diluent (ml)	CSPD (ml)
10	60	3
11	65	3.25
12	70	3.5
13	75	3.75
14	80	4
15	85	4.25
16	90	4.5
17	95	4.75
18	100	5
19	105	5.25
20	110	5.5
21	115	5.75
22	120	6
23	125	6.25
24	130	6.5
25	135	6.75

26	140	7
27	145	7.25
28	150	7.5
29	155	7.75
30	160	8
31	165	8.25
32	170	8.5
33	175	8.75
34	180	9
35	185	9.25
36	190	9.5
37	195	9.75
38	200	10
39	205	10.25
40	210	10.5
41	215	10.75
42	220	11
43	225	11.25
44	230	11.5
45	235	11.75
46	240	12
47	245	12.25
48	250	12.5
49	255	12.75
50	260	13

---

**Example 18: High-Throughput Screening Assay Identifying Changes in Small Molecule Concentration and Membrane Permeability**

Binding of a ligand to a receptor is known to alter intracellular levels of small molecules, such as calcium, potassium, sodium, and pH, as well as alter membrane potential. These alterations can be measured in an assay to identify supernatants which bind to receptors of a particular cell. Although the following protocol describes an assay for calcium, this protocol can easily be modified to detect changes in potassium, sodium, pH, membrane potential, or any other small molecule which is detectable by a fluorescent probe.

The following assay uses Fluorometric Imaging Plate Reader ("FLIPR") to measure changes in fluorescent molecules (Molecular Probes) that bind small



molecules. Clearly, any fluorescent molecule detecting a small molecule can be used instead of the calcium fluorescent molecule, fluo-4 (Molecular Probes, Inc.; catalog no. F-14202), used here.

For adherent cells, seed the cells at 10,000 -20,000 cells/well in a Co-star  
5 black 96-well plate with clear bottom. The plate is incubated in a CO<sub>2</sub> incubator for 20 hours. The adherent cells are washed two times in Biotek washer with 200 ul of HBSS (Hank's Balanced Salt Solution) leaving 100 ul of buffer after the final wash.

A stock solution of 1 mg/ml fluo-4 is made in 10% pluronic acid DMSO. To  
load the cells with fluo-4, 50 ul of 12 ug/ml fluo-4 is added to each well. The plate  
10 is incubated at 37 degrees C in a CO<sub>2</sub> incubator for 60 min. The plate is washed four times in the Biotek washer with HBSS leaving 100 ul of buffer.

For non-adherent cells, the cells are spun down from culture media. Cells are re-suspended to  $2-5 \times 10^6$  cells/ml with HBSS in a 50-ml conical tube. 4 ul of 1 mg/ml fluo-4 solution in 10% pluronic acid DMSO is added to each ml of cell suspension.  
15 The tube is then placed in a 37 degrees C water bath for 30-60 min. The cells are washed twice with HBSS, resuspended to  $1 \times 10^6$  cells/ml, and dispensed into a microplate, 100 ul/well. The plate is centrifuged at 1000 rpm for 5 min. The plate is then washed once in Denley CellWash with 200 ul, followed by an aspiration step to 100 ul final volume.

20 For a non-cell based assay, each well contains a fluorescent molecule, such as fluo-4. The supernatant is added to the well, and a change in fluorescence is detected.

To measure the fluorescence of intracellular calcium, the FLIPR is set for the following parameters: (1) System gain is 300-800 mW; (2) Exposure time is 0.4  
25 second; (3) Camera F/stop is F/2; (4) Excitation is 488 nm; (5) Emission is 530 nm; and (6) Sample addition is 50 ul. Increased emission at 530 nm indicates an

extracellular signaling event which has resulted in an increase in the intracellular  $\text{Ca}^{++}$  concentration.

**Example 19: High-Throughput Screening Assay Identifying Tyrosine Kinase**

**5    Activity**

The Protein Tyrosine Kinases (PTK) represent a diverse group of transmembrane and cytoplasmic kinases. Within the Receptor Protein Tyrosine Kinase (RPTK) group are receptors for a range of mitogenic and metabolic growth factors including the PDGF, FGF, EGF, NGF, HGF and Insulin receptor subfamilies. In addition there are a large family of RPTKs for which the corresponding ligand is unknown. Ligands for RPTKs include mainly secreted small proteins, but also membrane-bound and extracellular matrix proteins.

Activation of RPTK by ligands involves ligand-mediated receptor dimerization, resulting in transphosphorylation of the receptor subunits and activation of the cytoplasmic tyrosine kinases. The cytoplasmic tyrosine kinases include receptor associated tyrosine kinases of the src-family (e.g., src, yes, lck, lyn, fyn) and non-receptor linked and cytosolic protein tyrosine kinases, such as the Jak family, members of which mediate signal transduction triggered by the cytokine superfamily of receptors (e.g., the Interleukins, Interferons, GM-CSF, and Leptin).

20        Because of the wide range of known factors capable of stimulating tyrosine kinase activity, the identification of novel human secreted proteins capable of activating tyrosine kinase signal transduction pathways are of interest. Therefore, the following protocol is designed to identify those novel human secreted proteins capable of activating the tyrosine kinase signal transduction pathways.

25        Seed target cells (e.g., primary keratinocytes) at a density of approximately 25,000 cells per well in a 96 well Loprodyne Silent Screen Plates purchased from

Nalge Nunc (Naperville, IL). The plates are sterilized with two 30 minute rinses with 100% ethanol, rinsed with water and dried overnight. Some plates are coated for 2 hr with 100 ml of cell culture grade type I collagen (50 mg/ml), gelatin (2%) or polylysine (50 mg/ml), all of which can be purchased from Sigma Chemicals (St. Louis, MO) or 10% Matrigel purchased from Becton Dickinson (Bedford, MA), or calf serum, rinsed with PBS and stored at 4 degree C. Cell growth on these plates is assayed by seeding 5,000 cells/well in growth medium and indirect quantitation of cell number through use of alamarBlue as described by the manufacturer Alamar Biosciences, Inc. (Sacramento, CA) after 48 hr. Falcon plate covers #3071 from Becton Dickinson (Bedford, MA) are used to cover the Loprodyne Silent Screen Plates. Falcon Microtest III cell culture plates can also be used in some proliferation experiments.

To prepare extracts, A431 cells are seeded onto the nylon membranes of Loprodyne plates (20,000/200ml/well) and cultured overnight in complete medium. Cells are quiesced by incubation in serum-free basal medium for 24 hr. After 5-20 minutes treatment with EGF (60ng/ml) or 50 ul of the supernatant produced in Example 11, the medium was removed and 100 ml of extraction buffer ((20 mM HEPES pH 7.5, 0.15 M NaCl, 1% Triton X-100, 0.1% SDS, 2 mM Na<sub>3</sub>VO<sub>4</sub>, 2 mM Na<sub>4</sub>P<sub>2</sub>O<sub>7</sub> and a cocktail of protease inhibitors (# 1836170) obtained from Boehringer Mannheim (Indianapolis, IN) is added to each well and the plate is shaken on a rotating shaker for 5 minutes at 4 degrees C. The plate is then placed in a vacuum transfer manifold and the extract filtered through the 0.45 mm membrane bottoms of each well using house vacuum. Extracts are collected in a 96-well catch/assay plate in the bottom of the vacuum manifold and immediately placed on ice. To obtain extracts clarified by centrifugation, the content of each well, after

detergent solubilization for 5 minutes, is removed and centrifuged for 15 minutes at 4 degrees C at 16,000 x g.

Test the filtered extracts for levels of tyrosine kinase activity. Although many methods of detecting tyrosine kinase activity are known, one method is described  
5 here.

Generally, the tyrosine kinase activity of a supernatant is evaluated by determining its ability to phosphorylate a tyrosine residue on a specific substrate (a biotinylated peptide). Biotinylated peptides that can be used for this purpose include PSK1 (corresponding to amino acids 6-20 of the cell division kinase cdc2-p34) and  
10 PSK2 (corresponding to amino acids 1-17 of gastrin). Both peptides are substrates for a range of tyrosine kinases and are available from Boehringer Mannheim.

The tyrosine kinase reaction is set up by adding the following components in order. First, add 10ul of 5uM Biotinylated Peptide, then 10ul ATP/Mg<sub>2+</sub> (5mM ATP/50mM MgCl<sub>2</sub>), then 10ul of 5x Assay Buffer (40mM imidazole hydrochloride,  
15 pH7.3, 40 mM beta-glycerophosphate, 1mM EGTA, 100mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 0.5 mg/ml BSA), then 5ul of Sodium Vanadate(1mM), and then 5ul of water. Mix the components gently and preincubate the reaction mix at 30 degrees C for 2 min. Initial the reaction by adding 10ul of the control enzyme or the filtered supernatant.

The tyrosine kinase assay reaction is then terminated by adding 10 ul of  
20 120mM EDTA and place the reactions on ice.

Tyrosine kinase activity is determined by transferring 50 ul aliquot of reaction mixture to a microtiter plate (MTP) module and incubating at 37 degrees C for 20 min. This allows the streptavidin coated 96 well plate to associate with the biotinylated peptide. Wash the MTP module with 300ul/well of PBS four times.  
25 Next add 75 ul of anti-phosphotyrosine antibody conjugated to horse radish

peroxidase(anti-P-Tyr-POD(0.5u/ml)) to each well and incubate at 37 degrees C for one hour. Wash the well as above.

Next add 100ul of peroxidase substrate solution (Boehringer Mannheim) and incubate at room temperature for at least 5 mins (up to 30 min). Measure the  
5 absorbance of the sample at 405 nm by using ELISA reader. The level of bound peroxidase activity is quantitated using an ELISA reader and reflects the level of tyrosine kinase activity.

#### **Example 20: High-Throughput Screening Assay Identifying Phosphorylation**

##### **10 Activity**

As a potential alternative and/or compliment to the assay of protein tyrosine kinase activity described in Example 19, an assay which detects activation (phosphorylation) of major intracellular signal transduction intermediates can also be used. For example, as described below one particular assay can detect tyrosine  
15 phosphorylation of the Erk-1 and Erk-2 kinases. However, phosphorylation of other molecules, such as Raf, JNK, p38 MAP, Map kinase kinase (MEK), MEK kinase, Src, Muscle specific kinase (MuSK), IRAK, Tec, and Janus, as well as any other phosphoserine, phosphotyrosine, or phosphothreonine molecule, can be detected by substituting these molecules for Erk-1 or Erk-2 in the following assay.

20 Specifically, assay plates are made by coating the wells of a 96-well ELISA plate with 0.1ml of protein G (1ug/ml) for 2 hr at room temp, (RT). The plates are then rinsed with PBS and blocked with 3% BSA/PBS for 1 hr at RT. The protein G plates are then treated with 2 commercial monoclonal antibodies (100ng/well) against Erk-1 and Erk-2 (1 hr at RT) (Santa Cruz Biotechnology). (To detect other molecules,  
25 this step can easily be modified by substituting a monoclonal antibody detecting any

of the above described molecules.) After 3-5 rinses with PBS, the plates are stored at 4 degrees C until use.

A431 cells are seeded at 20,000/well in a 96-well Loprodyne filterplate and cultured overnight in growth medium. The cells are then starved for 48 hr in basal  
5 medium (DMEM) and then treated with EGF (6ng/well) or 50 ul of the supernatants obtained in Example 11 for 5-20 minutes. The cells are then solubilized and extracts filtered directly into the assay plate.

After incubation with the extract for 1 hr at RT, the wells are again rinsed. As a positive control, a commercial preparation of MAP kinase (10ng/well) is used in  
10 place of A431 extract. Plates are then treated with a commercial polyclonal (rabbit) antibody (1ug/ml) which specifically recognizes the phosphorylated epitope of the Erk-1 and Erk-2 kinases (1 hr at RT). This antibody is biotinylated by standard procedures. The bound polyclonal antibody is then quantitated by successive incubations with Europium-streptavidin and Europium fluorescence enhancing  
15 reagent in the Wallac DELFIA instrument (time-resolved fluorescence). An increased fluorescent signal over background indicates a phosphorylation.

**Example 21: Method of Determining Alterations in a Gene Corresponding to a Polynucleotide**

20 RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:X. Suggested PCR conditions consist of 35 cycles at 95 degrees C for  
25 30 seconds; 60-120 seconds at 52-58 degrees C; and 60-120 seconds at 70 degrees C, using buffer solutions described in Sidransky et al., Science 252:706 (1991).

PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

PCR products is cloned into T-tailed vectors as described in Holton et al., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 2 are nick-translated with digoxigenindeoxy-uridine 5'-triphosphate (Boehringer Mannheim), and FISH performed as described in Johnson et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, VT) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, AZ) and variable excitation wavelength filters. (Johnson et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, NC.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and

translocations. These alterations are used as a diagnostic marker for an associated disease.

**Example 22: Method of Detecting Abnormal Levels of a Polypeptide in a**

**5 Biological Sample**

A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their  
10 particular needs.

For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described in  
15 Example 10. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

The coated wells are then incubated for > 2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled  
20 water to remove unbounded polypeptide.

Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

25 Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room



temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard  
5 curve.

### **Example 23: Formulation**

The invention also provides methods of treatment and/or prevention diseases, disorders, and/or conditions (such as, for example, any one or more of the diseases or  
10 disorders disclosed herein) by administration to a subject of an effective amount of a Therapeutic. By therapeutic is meant a polynucleotides or polypeptides of the invention (including fragments and variants), agonists or antagonists thereof, and/or antibodies thereto, in combination with a pharmaceutically acceptable carrier type (e.g., a sterile carrier).

15 The Therapeutic will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the Therapeutic alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus  
20 determined by such considerations.

As a general proposition, the total pharmaceutically effective amount of the Therapeutic administered parenterally per dose will be in the range of about 1 µg/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01  
25 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the Therapeutic is typically administered at a

dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears  
5 to vary depending on the desired effect.

Therapeutics can be administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray.

"Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid  
10 filler, diluent, encapsulating material or formulation auxiliary of any. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Therapeutics of the invention are also suitably administered by sustained-  
15 release systems. Suitable examples of sustained-release Therapeutics are administered orally, rectally, parenterally, intracistemally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), buccally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or  
20 formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

Therapeutics of the invention are also suitably administered by sustained-  
release systems. Suitable examples of sustained-release Therapeutics include suitable  
25 polymeric materials (such as, for example, semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules), suitable hydrophobic materials

(for example as an emulsion in an acceptable oil) or ion exchange resins, and sparingly soluble derivatives (such as, for example, a sparingly soluble salt).

Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al.,  
5 Biopolymers 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (Langer et al., Id.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988).

Sustained-release Therapeutics also include liposomally entrapped  
10 Therapeutics of the invention (*see* generally, Langer, *Science* 249:1527-1533 (1990); Treat et al., in *Liposomes in the Therapy of Infectious Disease and Cancer*, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 317 -327 and 353-365 (1989)). Liposomes containing the Therapeutic are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. (USA) 82:3688-3692 (1985); Hwang  
15 et al., Proc. Natl. Acad. Sci.(USA) 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the  
20 optimal Therapeutic.

In yet an additional embodiment, the Therapeutics of the invention are delivered by way of a pump (*see* Langer, *supra*; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)).

25 Other controlled release systems are discussed in the review by Langer (*Science* 249:1527-1533 (1990)).

For parenteral administration, in one embodiment, the Therapeutic is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to the Therapeutic.

Generally, the formulations are prepared by contacting the Therapeutic uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

The Therapeutic is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

5           Any pharmaceutical used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutics generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

10           Therapeutics ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous Therapeutic solution, and the resulting mixture is lyophilized. The infusion solution is prepared by  
15   reconstituting the lyophilized Therapeutic using bacteriostatic Water-for-Injection.

          The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the Therapeutics of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or  
20   biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the Therapeutics may be employed in conjunction with other therapeutic compounds.

          The Therapeutics of the invention may be administered alone or in combination with adjuvants. Adjuvants that may be administered with the  
25   Therapeutics of the invention include, but are not limited to, alum, alum plus deoxycholate (ImmunoAg), MTP-PE (Biocine Corp.), QS21 (Genentech, Inc.), BCG,

and MPL. In a specific embodiment, Therapeutics of the invention are administered in combination with alum. In another specific embodiment, Therapeutics of the invention are administered in combination with QS-21. Further adjuvants that may be administered with the Therapeutics of the invention include, but are not limited to,

5 Monophosphoryl lipid immunomodulator, AdjuVax 100a, QS-21, QS-18, CRL1005, Aluminum salts, MF-59, and Virosomal adjuvant technology. Vaccines that may be administered with the Therapeutics of the invention include, but are not limited to, vaccines directed toward protection against MMR (measles, mumps, rubella), polio, varicella, tetanus/diphtheria, hepatitis A, hepatitis B, haemophilus influenzae B,

10 whooping cough, pneumonia, influenza, Lyme's Disease, rotavirus, cholera, yellow fever, Japanese encephalitis, poliomyelitis, rabies, typhoid fever, and pertussis. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations in which the combined agents are administered together as a therapeutic

15 mixture, and also procedures in which the combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given first, followed by the second.

The Therapeutics of the invention may be administered alone or in

20 combination with other therapeutic agents. Therapeutic agents that may be administered in combination with the Therapeutics of the invention, include but not limited to, other members of the TNF family, chemotherapeutic agents, antibiotics, steroidal and non-steroidal anti-inflammatories, conventional immunotherapeutic agents, cytokines and/or growth factors. Combinations may be administered either

25 concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations in which the combined agents are

administered together as a therapeutic mixture, and also procedures in which the combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given  
5 first, followed by the second.

In one embodiment, the Therapeutics of the invention are administered in combination with members of the TNF family. TNF, TNF-related or TNF-like molecules that may be administered with the Therapeutics of the invention include, but are not limited to, soluble forms of TNF-alpha, lymphotoxin-alpha (LT-alpha, also known as TNF-beta), LT-beta (found in complex heterotrimer LT-alpha2-beta),  
10 OPGL, FasL, CD27L, CD30L, CD40L, 4-1BBL, DcR3, OX40L, TNF-gamma (International Publication No. WO 96/14328), AIM-I (International Publication No. WO 97/33899), endokine-alpha (International Publication No. WO 98/07880), TR6 (International Publication No. WO 98/30694), OPG, and neutrokin-alpha  
15 (International Publication No. WO 98/18921, OX40, and nerve growth factor (NGF), and soluble forms of Fas, CD30, CD27, CD40 and 4-1BB, TR2 (International Publication No. WO 96/34095), DR3 (International Publication No. WO 97/33904), DR4 (International Publication No. WO 98/32856), TR5 (International Publication No. WO 98/30693), TR6 (International Publication No. WO 98/30694), TR7  
20 (International Publication No. WO 98/41629), TRANK, TR9 (International Publication No. WO 98/56892), TR10 (International Publication No. WO 98/54202), 312C2 (International Publication No. WO 98/06842), and TR12, and soluble forms CD154, CD70, and CD153.

In certain embodiments, Therapeutics of the invention are administered in  
25 combination with antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or protease inhibitors.

Nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, RETROVIR™ (zidovudine/AZT), VIDEX™ (didanosine/ddI), HIVID™ (zalcitabine/ddC), ZERIT™ (stavudine/d4T), EPIVIR™ (lamivudine/3TC), and COMBIVIR™ (zidovudine/lamivudine). Non-nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, VIRAMUNE™ (nevirapine), RESCRIPTOR™ (delavirdine), and SUSTIVA™ (efavirenz). Protease inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, CRIXIVAN™ (indinavir), NORVIR™ (ritonavir), INVIRASE™ (saquinavir), and VIRACEPT™ (nelfinavir). In a specific embodiment, antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or protease inhibitors may be used in any combination with Therapeutics of the invention to treat AIDS and/or to prevent or treat HIV infection.

15 In other embodiments, Therapeutics of the invention may be administered in combination with anti-opportunistic infection agents. Anti-opportunistic agents that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, TRIMETHOPRIM-SULFAMETHOXAZOLE™, DAPSONE™, PENTAMIDINE™, ATOVAQUONE™, ISONIAZID™, 20 RIFAMPIN™, PYRAZINAMIDE™, ETHAMBUTOL™, RIFABUTIN™, CLARITHROMYCIN™, AZITHROMYCIN™, GANCICLOVIR™, FOSCARNET™, CIDOFOVIR™, FLUCONAZOLE™, ITRACONAZOLE™, KETOCONAZOLE™, ACYCLOVIR™, FAMCICOLVIR™, PYRIMETHAMINE™, LEUCOVORIN™, NEUPOGEN™ (filgrastim/G-CSF), and LEUKINE™ 25 (sargramostim/GM-CSF). In a specific embodiment, Therapeutics of the invention are used in any combination with TRIMETHOPRIM-SULFAMETHOXAZOLE™,



DAPSONE™, PENTAMIDINE™, and/or ATOVAQUONE™ to prophylactically treat or prevent an opportunistic *Pneumocystis carinii* pneumonia infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ISONIAZID™, RIFAMPIN™, PYRAZINAMIDE™, and/or

5 ETHAMBUTOL™ to prophylactically treat or prevent an opportunistic *Mycobacterium avium* complex infection. In another specific embodiment, Therapeutics of the invention are used in any combination with RIFABUTIN™, CLARITHROMYCIN™, and/or AZITHROMYCIN™ to prophylactically treat or prevent an opportunistic *Mycobacterium tuberculosis* infection. In another specific

10 embodiment, Therapeutics of the invention are used in any combination with GANCICLOVIR™, FOSCARNET™, and/or CIDOFOVIR™ to prophylactically treat or prevent an opportunistic cytomegalovirus infection. In another specific embodiment, Therapeutics of the invention are used in any combination with FLUCONAZOLE™, ITRACONAZOLE™, and/or KETOCONAZOLE™ to

15 prophylactically treat or prevent an opportunistic fungal infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ACYCLOVIR™ and/or FAMCICOLVIR™ to prophylactically treat or prevent an opportunistic herpes simplex virus type I and/or type II infection. In another specific embodiment, Therapeutics of the invention are used in any combination with

20 PYRIMETHAMINE™ and/or LEUCOVORIN™ to prophylactically treat or prevent an opportunistic *Toxoplasma gondii* infection. In another specific embodiment, Therapeutics of the invention are used in any combination with LEUCOVORIN™ and/or NEUPOGEN™ to prophylactically treat or prevent an opportunistic bacterial infection.

25 In a further embodiment, the Therapeutics of the invention are administered in combination with an antiviral agent. Antiviral agents that may be administered

with the Therapeutics of the invention include, but are not limited to, acyclovir, ribavirin, amantadine, and remantidine.

In a further embodiment, the Therapeutics of the invention are administered in combination with an antibiotic agent. Antibiotic agents that may be administered  
5 with the Therapeutics of the invention include, but are not limited to, amoxicillin, beta-lactamases, aminoglycosides, beta-lactam (glycopeptide), beta-lactamases, Clindamycin, chloramphenicol, cephalosporins, ciprofloxacin, ciprofloxacin, erythromycin, fluoroquinolones, macrolides, metronidazole, penicillins, quinolones, rifampin, streptomycin, sulfonamide, tetracyclines, trimethoprim, trimethoprim-  
10 sulfamthoxazole, and vancomycin.

Conventional nonspecific immunosuppressive agents, that may be administered in combination with the Therapeutics of the invention include, but are not limited to, steroids, cyclosporine, cyclosporine analogs, cyclophosphamide methylprednisone, prednisone, azathioprine, FK-506, 15-deoxyspergualin, and other  
15 immunosuppressive agents that act by suppressing the function of responding T cells.

In specific embodiments, Therapeutics of the invention are administered in combination with immunosuppressants. Immunosuppressants preparations that may be administered with the Therapeutics of the invention include, but are not limited to, ORTHOCLONE™ (OKT3), SANDIMMUNE™/NEORAL™/SANGDYA™  
20 (cyclosporin), PROGRAF™ (tacrolimus), CELLCEPT™ (mycophenolate), Azathioprine, glucorticosteroids, and RAPAMUNE™ (sirolimus). In a specific embodiment, immunosuppressants may be used to prevent rejection of organ or bone marrow transplantation.

In an additional embodiment, Therapeutics of the invention are administered  
25 alone or in combination with one or more intravenous immune globulin preparations. Intravenous immune globulin preparations that may be administered with the

Therapeutics of the invention include, but not limited to, GAMMAR™, IVEEGAM™, SANDOGLOBULIN™, GAMMAGARD S/D™, and GAMIMUNE™.

In a specific embodiment, Therapeutics of the invention are administered in combination with intravenous immune globulin preparations in transplantation  
5 therapy (e.g., bone marrow transplant).

In an additional embodiment, the Therapeutics of the invention are administered alone or in combination with an anti-inflammatory agent. Anti-inflammatory agents that may be administered with the Therapeutics of the invention include, but are not limited to, glucocorticoids and the nonsteroidal anti-  
10 inflammatories, aminoarylcarboxylic acid derivatives, arylacetic acid derivatives, arylbutyric acid derivatives, arylcarboxylic acids, arylpropionic acid derivatives, pyrazoles, pyrazolones, salicylic acid derivatives, thiazinecarboxamides, e-acetamidocaproic acid, S-adenosylmethionine, 3-amino-4-hydroxybutyric acid, amixetrine, bendazac, benzydamine, bucolome, difenpiramide, ditazol, emorfazone,  
15 guaiazulene, nabumetone, nimesulide, orgotein, oxaceprol, paranyline, perisoxal, pifoxime, proquazone, proxazole, and tenidap.

In another embodiment, compositions of the invention are administered in combination with a chemotherapeutic agent. Chemotherapeutic agents that may be administered with the Therapeutics of the invention include, but are not limited to,  
20 antibiotic derivatives (e.g., doxorubicin, bleomycin, daunorubicin, and dactinomycin); antiestrogens (e.g., tamoxifen); antimetabolites (e.g., fluorouracil, 5-FU, methotrexate, floxuridine, interferon alpha-2b, glutamic acid, plicamycin, mercaptopurine, and 6-thioguanine); cytotoxic agents (e.g., carmustine, BCNU, lomustine, CCNU, cytosine arabinoside, cyclophosphamide, estramustine,  
25 hydroxyurea, procarbazine, mitomycin, busulfan, cis-platin, and vincristine sulfate); hormones (e.g., medroxyprogesterone, estramustine phosphate sodium, ethinyl

estradiol, estradiol, megestrol acetate, methyltestosterone, diethylstilbestrol diphosphate, chlorotrianisene, and testolactone); nitrogen mustard derivatives (e.g., mephalen, chorambucil, mechlorethamine (nitrogen mustard) and thiotepa); steroids and combinations (e.g., bethamethasone sodium phosphate); and others (e.g.,  
5 dicarbazine, asparaginase, mitotane, vincristine sulfate, vinblastine sulfate, and etoposide).

In a specific embodiment, Therapeutics of the invention are administered in combination with CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) or any combination of the components of CHOP. In another  
10 embodiment, Therapeutics of the invention are administered in combination with Rituximab. In a further embodiment, Therapeutics of the invention are administered with Rituxmab and CHOP, or Rituxmab and any combination of the components of CHOP.

In an additional embodiment, the Therapeutics of the invention are  
15 administered in combination with cytokines. Cytokines that may be administered with the Therapeutics of the invention include, but are not limited to, IL2, IL3, IL4, IL5, IL6, IL7, IL10, IL12, IL13, IL15, anti-CD40, CD40L, IFN-gamma and TNF-alpha. In another embodiment, Therapeutics of the invention may be administered with any interleukin, including, but not limited to, IL-1alpha, IL-1beta, IL-2, IL-3,  
20 IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, and IL-21.

In an additional embodiment, the Therapeutics of the invention are administered in combination with angiogenic proteins. Angiogenic proteins that may be administered with the Therapeutics of the invention include, but are not limited to,  
25 Glioma Derived Growth Factor (GDGF), as disclosed in European Patent Number EP-399816; Platelet Derived Growth Factor-A (PDGF-A), as disclosed in European

Patent Number EP-682110; Platelet Derived Growth Factor-B (PDGF-B), as disclosed in European Patent Number EP-282317; Placental Growth Factor (PlGF), as disclosed in International Publication Number WO 92/06194; Placental Growth Factor-2 (PlGF-2), as disclosed in Hauser et al., Growth Factors, 4:259-268 (1993);

5 Vascular Endothelial Growth Factor (VEGF), as disclosed in International Publication Number WO 90/13649; Vascular Endothelial Growth Factor-A (VEGF-A), as disclosed in European Patent Number EP-506477; Vascular Endothelial Growth Factor-2 (VEGF-2), as disclosed in International Publication Number WO 96/39515; Vascular Endothelial Growth Factor B (VEGF-3); Vascular Endothelial Growth

10 Factor B-186 (VEGF-B186), as disclosed in International Publication Number WO 96/26736; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/02543; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/07832; and Vascular Endothelial Growth Factor-E (VEGF-E), as disclosed in German Patent

15 Number DE19639601. The above mentioned references are incorporated herein by reference herein.

In an additional embodiment, the Therapeutics of the invention are administered in combination with hematopoietic growth factors. Hematopoietic growth factors that may be administered with the Therapeutics of the invention

20 include, but are not limited to, LEUKINE™ (SARGRAMOSTIM™) and NEUPOGEN™ (FILGRASTIM™).

In an additional embodiment, the Therapeutics of the invention are administered in combination with Fibroblast Growth Factors. Fibroblast Growth Factors that may be administered with the Therapeutics of the invention include, but

25 are not limited to, FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9, FGF-10, FGF-11, FGF-12, FGF-13, FGF-14, and FGF-15.

In additional embodiments, the Therapeutics of the invention are administered in combination with other therapeutic or prophylactic regimens, such as, for example, radiation therapy.

5    **Example 24: Method of Treating Decreased Levels of the Polypeptide**

          The present invention relates to a method for treating an individual in need of an increased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an agonist of the invention (including polypeptides of the  
10    invention). Moreover, it will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering  
15    to such an individual a Therapeutic comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

          For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on  
20    administration and formulation, are provided in Example 23.

**Example 25: Method of Treating Increased Levels of the Polypeptide**

          The present invention also relates to a method of treating an individual in need of a decreased level of a polypeptide of the invention in the body comprising  
25    administering to such an individual a composition comprising a therapeutically

effective amount of an antagonist of the invention (including polypeptides and antibodies of the invention).

In one example, antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided in Example 23.

#### **Example 26: Method of Treatment Using Gene Therapy-Ex Vivo**

One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37 degree C for approximately one week.

At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

pMV-7 (Kirschmeier, P.T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

5       The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 1 using primers and having appropriate restriction sites and initiation/stop codons, if necessary. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine  
10 sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene  
15 of interest properly inserted.

      The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector.  
20 The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

      Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove  
25 detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the



media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been  
5 efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

10 **Example 27: Gene Therapy Using Endogenous Genes Corresponding To Polynucleotides of the Invention**

Another method of gene therapy according to the present invention involves operably associating the endogenous polynucleotide sequence of the invention with a promoter via homologous recombination as described, for example, in U.S. Patent  
15 NO: 5,641,670, issued June 24, 1997; International Publication NO: WO 96/29411, published September 26, 1996; International Publication NO: WO 94/12650, published August 4, 1994; Koller et al., *Proc. Natl. Acad. Sci. USA*, 86:8932-8935 (1989); and Zijlstra et al., *Nature*, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not expressed in  
20 the cells, or is expressed at a lower level than desired.

Polynucleotide constructs are made which contain a promoter and targeting sequences, which are homologous to the 5' non-coding sequence of endogenous polynucleotide sequence, flanking the promoter. The targeting sequence will be sufficiently near the 5' end of the polynucleotide sequence so the promoter will be  
25 operably linked to the endogenous sequence upon homologous recombination. The promoter and the targeting sequences can be amplified using PCR. Preferably, the

amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified  
5 promoter.

The amplified promoter and the amplified targeting sequences are digested with the appropriate restriction enzymes and subsequently treated with calf intestinal phosphatase. The digested promoter and digested targeting sequences are added together in the presence of T4 DNA ligase. The resulting mixture is maintained under  
10 conditions appropriate for ligation of the two fragments. The construct is size fractionated on an agarose gel then purified by phenol extraction and ethanol precipitation.

In this Example, the polynucleotide constructs are administered as naked polynucleotides via electroporation. However, the polynucleotide constructs may also  
15 be administered with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, precipitating agents, etc. Such methods of delivery are known in the art.

Once the cells are transfected, homologous recombination will take place which results in the promoter being operably linked to the endogenous polynucleotide  
20 sequence. This results in the expression of polynucleotide corresponding to the polynucleotide in the cell. Expression may be detected by immunological staining, or any other method known in the art.

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in DMEM + 10% fetal calf serum. Exponentially growing or early stationary  
25 phase fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining

cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of electroporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl, 5 mM KCl, 0.7 mM Na<sub>2</sub> HPO<sub>4</sub>, 6 mM dextrose). The cells are recentrifuged, the supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension contains approximately 3X10<sup>6</sup> cells/ml. Electroporation should be performed immediately following resuspension.

Plasmid DNA is prepared according to standard techniques. For example, to construct a plasmid for targeting to the locus corresponding to the polynucleotide of the invention, plasmid pUC18 (MBI Fermentas, Amherst, NY) is digested with HindIII. The CMV promoter is amplified by PCR with an XbaI site on the 5' end and a BamHI site on the 3' end. Two non-coding sequences are amplified via PCR: one non-coding sequence (fragment 1) is amplified with a HindIII site at the 5' end and an XbaI site at the 3' end; the other non-coding sequence (fragment 2) is amplified with a BamHI site at the 5' end and a HindIII site at the 3' end. The CMV promoter and the fragments (1 and 2) are digested with the appropriate enzymes (CMV promoter - XbaI and BamHI; fragment 1 - XbaI; fragment 2 - BamHI) and ligated together. The resulting ligation product is digested with HindIII, and ligated with the HindIII-digested pUC18 plasmid.

Plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad). The final DNA concentration is generally at least 120 µg/ml. 0.5 ml of the cell suspension (containing approximately 1.5X10<sup>6</sup> cells) is then added to the cuvette, and the cell suspension and DNA solutions are gently mixed. Electroporation is performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at 960 µF and 250-300 V, respectively. As voltage increases, cell survival decreases, but the percentage of surviving cells that stably incorporate the introduced DNA into their

genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer  
5 pipette. The cells are added directly to 10 ml of prewarmed nutrient media (DMEM with 15% calf serum) in a 10 cm dish and incubated at 37 degree C. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a further 16-24 hours.

The engineered fibroblasts are then injected into the host, either alone or after  
10 having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product. The fibroblasts can then be introduced into a patient as described above.

**Example 28: Method of Treatment Using Gene Therapy - In Vivo**

15 Another aspect of the present invention is using *in vivo* gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter  
20 or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Patent NO. 5693622, 5705151, 5580859; Tabata et al., Cardiovasc. Res. 35(3):470-479 (1997); Chao et al., Pharmacol. Res. 35(6):517-522 (1997); Wolff, Neuromuscul. Disord. 7(5):314-318  
25 (1997); Schwartz et al., Gene Ther. 3(5):405-411 (1996); Tsurumi et al., Circulation 94(12):3281-3290 (1996) (incorporated herein by reference).

The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Felgner P.L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid,

mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. *In vivo* muscle cells are particularly competent in their ability to take up and express polynucleotides.

For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

The dose response effects of injected polynucleotide in muscle *in vivo* is

determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps  
5 muscles of mice are then injected with various amounts of the template DNA.

Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a 1 cc syringe through a 27 gauge needle over  
10 one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the  
15 individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice.  
20 The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

#### **Example 29: Transgenic Animals.**

25 The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters,

guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, *e.g.*, baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy  
5 protocol.

Any technique known in the art may be used to introduce the transgene (*i.e.*, polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994);  
10 Carver et al., Biotechnology (NY) 11:1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or  
15 embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, *e.g.*, Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pluripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of  
20 such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to  
25 quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).



The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, *i.e.*, mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, *e.g.*, head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis

of tissue samples obtained from the animal, *in situ* hybridization analysis, and reverse transcriptase-PCR (rt-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

5           Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher  
10   levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the  
15   transgene on a distinct background that is appropriate for an experimental model of interest.

Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions  
20   associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

#### **Example 30: Knock-Out Animals.**

Endogenous gene expression can also be reduced by inactivating or "knocking  
25   out" the gene and/or its promoter using targeted homologous recombination. (*E.g.*, see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-

512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention *in vivo*. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (*e.g.*, see Thomas & Capecchi 1987 and Thompson 1989, *supra*). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors that will be apparent to those of skill in the art.

In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (*e.g.*, knockouts) are administered to a patient *in vivo*. Such cells may be obtained from the patient (*i.e.*, animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (*e.g.*, lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered *in vitro* using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention,

e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Patent No. 5,399,349; and Mulligan & Wilson, U.S. Patent No. 5,460,959 each of which is incorporated by reference herein in its entirety).

When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

**Example 31: Production of an Antibody****a) Hybridoma Technology**

The antibodies of the present invention can be prepared by a variety of methods. (See, Current Protocols, Chapter 2.) As one example of such methods, cells expressing polypeptide(s) of the invention are administered to an animal to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of polypeptide(s) of the invention is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

Monoclonal antibodies specific for polypeptide(s) of the invention are prepared using hybridoma technology. (Kohler et al., Nature 256:495 (1975); Kohler et al., Eur. J. Immunol. 6:511 (1976); Kohler et al., Eur. J. Immunol. 6:292 (1976); Hammerling et al., in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981)). In general, an animal (preferably a mouse) is immunized with polypeptide(s) of the invention or, more preferably, with a secreted polypeptide-expressing cell. Such polypeptide-expressing cells are cultured in any suitable tissue culture medium, preferably in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56°C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 µg/ml of streptomycin.

The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the ATCC. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as

described by Wands et al. (Gastroenterology 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the polypeptide(s) of the invention.

Alternatively, additional antibodies capable of binding to polypeptide(s) of the invention can be produced in a two-step procedure using anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the protein-specific antibody can be blocked by polypeptide(s) of the invention. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and are used to immunize an animal to induce formation of further protein-specific antibodies.

For in vivo use of antibodies in humans, an antibody is "humanized". Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric and humanized antibodies are known in the art and are discussed herein. (See, for review, Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985).)

b) Isolation Of Antibody Fragments Directed Against  
Polypeptide(s) From A Library Of scFvs

Naturally occurring V-genes isolated from human PBLs are constructed into a library of antibody fragments which contain reactivities against polypeptide(s) of the invention to which the donor may or may not have been exposed (see e.g., U.S. Patent 5,885,793 incorporated herein by reference in its entirety).

5           Rescue of the Library.

A library of scFvs is constructed from the RNA of human PBLs as described in PCT publication WO 92/01047. To rescue phage displaying antibody fragments, approximately 10<sup>9</sup> E. coli harboring the phagemid are used to inoculate 50 ml of 2xTY containing 1% glucose and 100 µg/ml of ampicillin (2xTY-AMP-GLU) and  
10 grown to an O.D. of 0.8 with shaking. Five ml of this culture is used to inoculate 50 ml of 2xTY-AMP-GLU, 2 x 10<sup>8</sup> TU of delta gene 3 helper (M13 delta gene III, see PCT publication WO 92/01047) are added and the culture incubated at 37°C for 45 minutes without shaking and then at 37°C for 45 minutes with shaking. The culture is centrifuged at 4000 r.p.m. for 10 min. and the pellet resuspended in 2 liters of 2xTY  
15 containing 100 µg/ml ampicillin and 50 µg/ml kanamycin and grown overnight. Phage are prepared as described in PCT publication WO 92/01047.

M13 delta gene III is prepared as follows: M13 delta gene III helper phage does not encode gene III protein, hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13 delta gene III  
20 particles are made by growing the helper phage in cells harboring a pUC19 derivative supplying the wild type gene III protein during phage morphogenesis. The culture is incubated for 1 hour at 37° C without shaking and then for a further hour at 37°C with shaking. Cells are spun down (IEC-Centra 8,400 r.p.m. for 10 min), resuspended in 300 ml 2xTY broth containing 100 µg ampicillin/ml and 25 µg kanamycin/ml (2xTY-  
25 AMP-KAN) and grown overnight, shaking at 37°C. Phage particles are purified and concentrated from the culture medium by two PEG-precipitations (Sambrook et al.,

1990), resuspended in 2 ml PBS and passed through a 0.45  $\mu$ m filter (Minisart NML; Sartorius) to give a final concentration of approximately  $10^{13}$  transducing units/ml (ampicillin-resistant clones).

#### Panning of the Library.

5        Immunotubes (Nunc) are coated overnight in PBS with 4 ml of either 100  $\mu$ g/ml or 10  $\mu$ g/ml of a polypeptide of the present invention. Tubes are blocked with 2% Marvel-PBS for 2 hours at 37°C and then washed 3 times in PBS. Approximately  $10^{13}$  TU of phage is applied to the tube and incubated for 30 minutes at room temperature tumbling on an over and under turntable and then left to stand for another  
10    1.5 hours. Tubes are washed 10 times with PBS 0.1% Tween-20 and 10 times with PBS. Phage are eluted by adding 1 ml of 100 mM triethylamine and rotating 15 minutes on an under and over turntable after which the solution is immediately neutralized with 0.5 ml of 1.0M Tris-HCl, pH 7.4. Phage are then used to infect 10 ml of mid-log E. coli TG1 by incubating eluted phage with bacteria for 30 minutes at  
15    37°C. The E. coli are then plated on TYE plates containing 1% glucose and 100  $\mu$ g/ml ampicillin. The resulting bacterial library is then rescued with delta gene 3 helper phage as described above to prepare phage for a subsequent round of selection. This process is then repeated for a total of 4 rounds of affinity purification with tube-washing increased to 20 times with PBS, 0.1% Tween-20 and 20 times with PBS for  
20    rounds 3 and 4.

#### Characterization of Binders.

Eluted phage from the 3rd and 4th rounds of selection are used to infect E. coli HB 2151 and soluble scFv is produced (Marks, et al., 1991) from single colonies for assay. ELISAs are performed with microtitre plates coated with either 10 pg/ml of  
25    the polypeptide of the present invention in 50 mM bicarbonate pH 9.6. Clones positive in ELISA are further characterized by PCR fingerprinting (see, e.g., PCT



publication WO 92/01047) and then by sequencing. These ELISA positive clones may also be further characterized by techniques known in the art, such as, for example, epitope mapping, binding affinity, receptor signal transduction, ability to block or competitively inhibit antibody/antigen binding, and competitive agonistic or  
5 antagonistic activity.

**Example 32: Assays Detecting Stimulation or Inhibition of B cell Proliferation and Differentiation**

Generation of functional humoral immune responses requires both soluble and  
10 cognate signaling between B-lineage cells and their microenvironment. Signals may impart a positive stimulus that allows a B-lineage cell to continue its programmed development, or a negative stimulus that instructs the cell to arrest its current developmental pathway. To date, numerous stimulatory and inhibitory signals have been found to influence B cell responsiveness including IL-2, IL-4, IL-5, IL-6, IL-7, IL10, IL-  
15 13, IL-14 and IL-15. Interestingly, these signals are by themselves weak effectors but can, in combination with various co-stimulatory proteins, induce activation, proliferation, differentiation, homing, tolerance and death among B cell populations.

One of the best studied classes of B-cell co-stimulatory proteins is the TNF-superfamily. Within this family CD40, CD27, and CD30 along with their respective  
20 ligands CD154, CD70, and CD153 have been found to regulate a variety of immune responses. Assays which allow for the detection and/or observation of the proliferation and differentiation of these B-cell populations and their precursors are valuable tools in determining the effects various proteins may have on these B-cell populations in terms of proliferation and differentiation. Listed below are two assays designed to allow for the  
25 detection of the differentiation, proliferation, or inhibition of B-cell populations and their precursors.

In Vitro Assay- Purified polypeptides of the invention, or truncated forms thereof, is assessed for its ability to induce activation, proliferation, differentiation or inhibition and/or death in B-cell populations and their precursors. The activity of the polypeptides of the invention on purified human tonsillar B cells, measured qualitatively over the dose  
5 range from 0.1 to 10,000 ng/mL, is assessed in a standard B-lymphocyte co-stimulation assay in which purified tonsillar B cells are cultured in the presence of either formalin-fixed *Staphylococcus aureus* Cowan I (SAC) or immobilized anti-human IgM antibody as the priming agent. Second signals such as IL-2 and IL-15 synergize with SAC and IgM crosslinking to elicit B cell proliferation as measured by tritiated-thymidine incorporation.  
10 Novel synergizing agents can be readily identified using this assay. The assay involves isolating human tonsillar B cells by magnetic bead (MACS) depletion of CD3-positive cells. The resulting cell population is greater than 95% B cells as assessed by expression of CD45R(B220).

Various dilutions of each sample are placed into individual wells of a 96-well plate  
15 to which are added  $10^5$  B-cells suspended in culture medium (RPMI 1640 containing 10% FBS,  $5 \times 10^{-5}$ M 2ME, 100U/ml penicillin, 10ug/ml streptomycin, and  $10^{-5}$  dilution of SAC) in a total volume of 150ul. Proliferation or inhibition is quantitated by a 20h pulse (1uCi/well) with  $^3$ H-thymidine (6.7 Ci/mM) beginning 72h post factor addition. The positive and negative controls are IL2 and medium respectively.

20 In Vivo Assay- BALB/c mice are injected (i.p.) twice per day with buffer only, or 2 mg/Kg of a polypeptide of the invention, or truncated forms thereof. Mice receive this treatment for 4 consecutive days, at which time they are sacrificed and various tissues and serum collected for analyses. Comparison of H&E sections from normal spleens and spleens treated with polypeptides of the invention identify the results of the activity of the  
25 polypeptides on spleen cells, such as the diffusion of peri-arterial lymphatic sheaths, and/or significant increases in the nucleated cellularity of the red pulp regions, which may

indicate the activation of the differentiation and proliferation of B-cell populations. Immunohistochemical studies using a B cell marker, anti-CD45R(B220), are used to determine whether any physiological changes to splenic cells, such as splenic disorganization, are due to increased B-cell representation within loosely defined B-cell zones that infiltrate established T-cell regions.

Flow cytometric analyses of the spleens from mice treated with polypeptide is used to indicate whether the polypeptide specifically increases the proportion of ThB+, CD45R(B220)dull B cells over that which is observed in control mice.

Likewise, a predicted consequence of increased mature B-cell representation in vivo is a relative increase in serum Ig titers. Accordingly, serum IgM and IgA levels are compared between buffer and polypeptide-treated mice.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

### **Example 33: T Cell Proliferation Assay**

A CD3-induced proliferation assay is performed on PBMCs and is measured by the uptake of <sup>3</sup>H-thymidine. The assay is performed as follows. Ninety-six well plates are coated with 100 µl/well of mAb to CD3 (HIT3a, Pharmingen) or isotype-matched control mAb (B33.1) overnight at 4 degrees C (1 µg/ml in .05M bicarbonate buffer, pH 9.5), then washed three times with PBS. PBMC are isolated by F/H gradient centrifugation from human peripheral blood and added to quadruplicate wells (5 x 10<sup>4</sup>/well) of mAb coated plates in RPMI containing 10% FCS and P/S in the presence of varying concentrations of polypeptides of the invention (total volume 200 ul). Relevant protein buffer and medium alone are controls. After 48 hr. culture at 37 degrees C, plates are spun for 2 min. at 1000 rpm and 100 µl of supernatant is removed and stored -20 degrees C for measurement of

IL-2 (or other cytokines) if effect on proliferation is observed. Wells are supplemented with 100 ul of medium containing 0.5 uCi of  $^3\text{H}$ -thymidine and cultured at 37 degrees C for 18-24 hr. Wells are harvested and incorporation of  $^3\text{H}$ -thymidine used as a measure of proliferation. Anti-CD3 alone is the positive control for proliferation. IL-2 (100 U/ml) is  
5 also used as a control which enhances proliferation. Control antibody which does not induce proliferation of T cells is used as the negative controls for the effects of polypeptides of the invention.

The studies described in this example tested activity of polypeptides of the invention. However, one skilled in the art could easily modify the exemplified studies to  
10 test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

**Example 34: Effect of Polypeptides of the Invention on the Expression of MHC Class II, Costimulatory and Adhesion Molecules and Cell Differentiation of**  
15 **Monocytes and Monocyte-Derived Human Dendritic Cells**

Dendritic cells are generated by the expansion of proliferating precursors found in the peripheral blood: adherent PBMC or elutriated monocytic fractions are cultured for 7-10 days with GM-CSF (50 ng/ml) and IL-4 (20 ng/ml). These dendritic cells have the characteristic phenotype of immature cells (expression of CD1, CD80, CD86, CD40 and  
20 MHC class II antigens). Treatment with activating factors, such as TNF- $\alpha$ , causes a rapid change in surface phenotype (increased expression of MHC class I and II, costimulatory and adhesion molecules, downregulation of FCyRII, upregulation of CD83). These changes correlate with increased antigen-presenting capacity and with functional maturation of the dendritic cells.

25 FACS analysis of surface antigens is performed as follows. Cells are treated 1-3 days with increasing concentrations of polypeptides of the invention or LPS (positive

control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC- or PE-labeled monoclonal antibodies for 30 minutes at 4 degrees C. After an additional wash, the labeled cells are analyzed by flow cytometry on a FACScan (Becton Dickinson).

5

Effect on the production of cytokines. Cytokines generated by dendritic cells, in particular IL-12, are important in the initiation of T-cell dependent immune responses. IL-12 strongly influences the development of Th1 helper T-cell immune response, and induces cytotoxic T and NK cell function. An ELISA is used to measure the IL-12 release as follows. Dendritic cells ( $10^6$ /ml) are treated with increasing concentrations of polypeptides of the invention for 24 hours. LPS (100 ng/ml) is added to the cell culture as positive control. Supernatants from the cell cultures are then collected and analyzed for IL-12 content using commercial ELISA kit (e.g., R & D Systems (Minneapolis, MN)). The standard protocols provided with the kits are used.

15

Effect on the expression of MHC Class II, costimulatory and adhesion molecules. Three major families of cell surface antigens can be identified on monocytes: adhesion molecules, molecules involved in antigen presentation, and Fc receptor. Modulation of the expression of MHC class II antigens and other costimulatory molecules, such as B7 and ICAM-1, may result in changes in the antigen presenting capacity of monocytes and ability to induce T cell activation. Increase expression of Fc receptors may correlate with improved monocyte cytotoxic activity, cytokine release and phagocytosis.

20

FACS analysis is used to examine the surface antigens as follows. Monocytes are treated 1-5 days with increasing concentrations of polypeptides of the invention or LPS (positive control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC- or PE-labeled monoclonal

25

antibodies for 30 minutes at 4 degreesC. After an additional wash, the labeled cells are analyzed by flow cytometry on a FACScan (Becton Dickinson).

Monocyte activation and/or increased survival. Assays for molecules that activate  
5 (or alternatively, inactivate) monocytes and/or increase monocyte survival (or  
alternatively, decrease monocyte survival) are known in the art and may routinely be  
applied to determine whether a molecule of the invention functions as an inhibitor or  
activator of monocytes. Polypeptides, agonists, or antagonists of the invention can be  
screened using the three assays described below. For each of these assays, Peripheral  
10 blood mononuclear cells (PBMC) are purified from single donor leukopacks (American  
Red Cross, Baltimore, MD) by centrifugation through a Histopaque gradient (Sigma).  
Monocytes are isolated from PBMC by counterflow centrifugal elutriation.

Monocyte Survival Assay. Human peripheral blood monocytes progressively lose  
15 viability when cultured in absence of serum or other stimuli. Their death results from  
internally regulated process (apoptosis). Addition to the culture of activating factors, such  
as TNF-alpha dramatically improves cell survival and prevents DNA fragmentation.  
Propidium iodide (PI) staining is used to measure apoptosis as follows. Monocytes are  
cultured for 48 hours in polypropylene tubes in serum-free medium (positive control), in  
20 the presence of 100 ng/ml TNF-alpha (negative control), and in the presence of varying  
concentrations of the compound to be tested. Cells are suspended at a concentration of  $2 \times 10^6$ /ml in PBS containing PI at a final concentration of 5  $\mu$ g/ml, and then incubated at room  
temperature for 5 minutes before FACScan analysis. PI uptake has been demonstrated to  
correlate with DNA fragmentation in this experimental paradigm.

Effect on cytokine release. An important function of monocytes/macrophages is their regulatory activity on other cellular populations of the immune system through the release of cytokines after stimulation. An ELISA to measure cytokine release is performed as follows. Human monocytes are incubated at a density of  $5 \times 10^5$  cells/ml with increasing concentrations of the a polypeptide of the invention and under the same conditions, but in the absence of the polypeptide. For IL-12 production, the cells are primed overnight with IFN (100 U/ml) in presence of a polypeptide of the invention. LPS (10 ng/ml) is then added. Conditioned media are collected after 24h and kept frozen until use. Measurement of TNF-alpha, IL-10, MCP-1 and IL-8 is then performed using a commercially available ELISA kit (e.g, R & D Systems (Minneapolis, MN)) and applying the standard protocols provided with the kit.

Oxidative burst. Purified monocytes are plated in 96-w plate at  $2 \times 10^5$  cell/well. Increasing concentrations of polypeptides of the invention are added to the wells in a total volume of 0.2 ml culture medium (RPMI 1640 + 10% FCS, glutamine and antibiotics). After 3 days incubation, the plates are centrifuged and the medium is removed from the wells. To the macrophage monolayers, 0.2 ml per well of phenol red solution (140 mM NaCl, 10 mM potassium phosphate buffer pH 7.0, 5.5 mM dextrose, 0.56 mM phenol red and 19 U/ml of HRPO) is added, together with the stimulant (200 nM PMA). The plates are incubated at 37°C for 2 hours and the reaction is stopped by adding 20  $\mu$ l 1N NaOH per well. The absorbance is read at 610 nm. To calculate the amount of  $H_2O_2$  produced by the macrophages, a standard curve of a  $H_2O_2$  solution of known molarity is performed for each experiment.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to

test the activity of polypeptides, polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

### **Example 35: Biological Effects of Polypeptides of the Invention**

#### **5 Astrocyte and Neuronal Assays**

Recombinant polypeptides of the invention, expressed in *Escherichia coli* and purified as described above, can be tested for activity in promoting the survival, neurite outgrowth, or phenotypic differentiation of cortical neuronal cells and for inducing the proliferation of glial fibrillary acidic protein immunopositive cells, astrocytes. The selection of cortical cells for the bioassay is based on the prevalent expression of FGF-1 and FGF-2 in cortical structures and on the previously reported enhancement of cortical neuronal survival resulting from FGF-2 treatment. A thymidine incorporation assay, for example, can be used to elucidate a polypeptide of the invention's activity on these cells.

Moreover, previous reports describing the biological effects of FGF-2 (basic FGF) on cortical or hippocampal neurons *in vitro* have demonstrated increases in both neuron survival and neurite outgrowth (Walicke et al., "Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension." *Proc. Natl. Acad. Sci. USA* 83:3012-3016. (1986), assay herein incorporated by reference in its entirety). However, reports from experiments done on PC-12 cells suggest that these two responses are not necessarily synonymous and may depend on not only which FGF is being tested but also on which receptor(s) are expressed on the target cells. Using the primary cortical neuronal culture paradigm, the ability of a polypeptide of the invention to induce neurite outgrowth can be compared to the response achieved with FGF-2 using, for example, a thymidine incorporation assay.

25

#### **Fibroblast and endothelial cell assays**



Human lung fibroblasts are obtained from Clonetics (San Diego, CA) and maintained in growth media from Clonetics. Dermal microvascular endothelial cells are obtained from Cell Applications (San Diego, CA). For proliferation assays, the human lung fibroblasts and dermal microvascular endothelial cells can be cultured at 5,000 cells/well in a 96-well plate for one day in growth medium. The cells are then incubated for one day in 0.1% BSA basal medium. After replacing the medium with fresh 0.1% BSA medium, the cells are incubated with the test proteins for 3 days. Alamar Blue (Alamar Biosciences, Sacramento, CA) is added to each well to a final concentration of 10%. The cells are incubated for 4 hr. Cell viability is measured by reading in a CytoFluor fluorescence reader. For the PGE<sub>2</sub> assays, the human lung fibroblasts are cultured at 5,000 cells/well in a 96-well plate for one day. After a medium change to 0.1% BSA basal medium, the cells are incubated with FGF-2 or polypeptides of the invention with or without IL-1 $\alpha$  for 24 hours. The supernatants are collected and assayed for PGE<sub>2</sub> by EIA kit (Cayman, Ann Arbor, MI). For the IL-6 assays, the human lung fibroblasts are cultured at 5,000 cells/well in a 96-well plate for one day. After a medium change to 0.1% BSA basal medium, the cells are incubated with FGF-2 or with or without polypeptides of the invention IL-1 $\alpha$  for 24 hours. The supernatants are collected and assayed for IL-6 by ELISA kit (Endogen, Cambridge, MA).

Human lung fibroblasts are cultured with FGF-2 or polypeptides of the invention for 3 days in basal medium before the addition of Alamar Blue to assess effects on growth of the fibroblasts. FGF-2 should show a stimulation at 10 - 2500 ng/ml which can be used to compare stimulation with polypeptides of the invention.

#### Parkinson Models.

The loss of motor function in Parkinson's disease is attributed to a deficiency of striatal dopamine resulting from the degeneration of the nigrostriatal dopaminergic

projection neurons. An animal model for Parkinson's that has been extensively characterized involves the systemic administration of 1-methyl-4 phenyl 1,2,3,6-tetrahydropyridine (MPTP). In the CNS, MPTP is taken-up by astrocytes and catabolized by monoamine oxidase B to 1-methyl-4-phenyl pyridine (MPP<sup>+</sup>) and released.

5 Subsequently, MPP<sup>+</sup> is actively accumulated in dopaminergic neurons by the high-affinity reuptake transporter for dopamine. MPP<sup>+</sup> is then concentrated in mitochondria by the electrochemical gradient and selectively inhibits nicotinamide adenine disphosphate: ubiquinone oxidoreductionase (complex I), thereby interfering with electron transport and eventually generating oxygen radicals.

10 It has been demonstrated in tissue culture paradigms that FGF-2 (basic FGF) has trophic activity towards nigral dopaminergic neurons (Ferrari et al., Dev. Biol. 1989). Recently, Dr. Unsicker's group has demonstrated that administering FGF-2 in gel foam implants in the striatum results in the near complete protection of nigral dopaminergic neurons from the toxicity associated with MPTP exposure (Otto and Unsicker, J.  
15 Neuroscience, 1990).

Based on the data with FGF-2, polypeptides of the invention can be evaluated to determine whether it has an action similar to that of FGF-2 in enhancing dopaminergic neuronal survival *in vitro* and it can also be tested *in vivo* for protection of dopaminergic neurons in the striatum from the damage associated with MPTP treatment. The potential  
20 effect of a polypeptide of the invention is first examined *in vitro* in a dopaminergic neuronal cell culture paradigm. The cultures are prepared by dissecting the midbrain floor plate from gestation day 14 Wistar rat embryos. The tissue is dissociated with trypsin and seeded at a density of 200,000 cells/cm<sup>2</sup> on polyorthinine-laminin coated glass coverslips. The cells are maintained in Dulbecco's Modified Eagle's medium and F12 medium  
25 containing hormonal supplements (N1). The cultures are fixed with paraformaldehyde after 8 days *in vitro* and are processed for tyrosine hydroxylase, a specific marker for

dopaminergic neurons, immunohistochemical staining. Dissociated cell cultures are prepared from embryonic rats. The culture medium is changed every third day and the factors are also added at that time.

Since the dopaminergic neurons are isolated from animals at gestation day 14, a developmental time which is past the stage when the dopaminergic precursor cells are proliferating, an increase in the number of tyrosine hydroxylase immunopositive neurons would represent an increase in the number of dopaminergic neurons surviving *in vitro*. Therefore, if a polypeptide of the invention acts to prolong the survival of dopaminergic neurons, it would suggest that the polypeptide may be involved in Parkinson's Disease.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 36: The Effect of Polypeptides of the Invention on the Growth of Vascular Endothelial Cells**

On day 1, human umbilical vein endothelial cells (HUVEC) are seeded at  $2-5 \times 10^4$  cells/35 mm dish density in M199 medium containing 4% fetal bovine serum (FBS), 16 units/ml heparin, and 50 units/ml endothelial cell growth supplements (ECGS, Biotechnology, Inc.). On day 2, the medium is replaced with M199 containing 10% FBS, 8 units/ml heparin. A polypeptide having the amino acid sequence of SEQ ID NO:Y, and positive controls, such as VEGF and basic FGF (bFGF) are added, at varying concentrations. On days 4 and 6, the medium is replaced. On day 8, cell number is determined with a Coulter Counter.

An increase in the number of HUVEC cells indicates that the polypeptide of the invention may proliferate vascular endothelial cells.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

5

**Example 37: Stimulatory Effect of Polypeptides of the Invention on the Proliferation of Vascular Endothelial Cells**

For evaluation of mitogenic activity of growth factors, the colorimetric MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)2H-tetrazolium) assay with the electron coupling reagent PMS (phenazine methosulfate) was  
10 performed (CellTiter 96 AQ, Promega). Cells are seeded in a 96-well plate (5,000 cells/well) in 0.1 mL serum-supplemented medium and are allowed to attach overnight. After serum-starvation for 12 hours in 0.5% FBS, conditions (bFGF, VEGF<sub>165</sub> or a polypeptide of the invention in 0.5% FBS) with or without Heparin (8 U/ml) are added to  
15 wells for 48 hours. 20 mg of MTS/PMS mixture (1:0.05) are added per well and allowed to incubate for 1 hour at 37°C before measuring the absorbance at 490 nm in an ELISA plate reader. Background absorbance from control wells (some media, no cells) is subtracted, and seven wells are performed in parallel for each condition. See, Leak *et al.* *In Vitro Cell. Dev. Biol.* 30A:512-518 (1994).

20 The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 38: Inhibition of PDGF-induced Vascular Smooth Muscle Cell****Proliferation Stimulatory Effect**

HAoSMC proliferation can be measured, for example, by BrdUrd incorporation.

5 Briefly, subconfluent, quiescent cells grown on the 4-chamber slides are transfected with CRP or FITC-labeled AT2-3LP. Then, the cells are pulsed with 10% calf serum and 6 mg/ml BrdUrd. After 24 h, immunocytochemistry is performed by using BrdUrd Staining Kit (Zymed Laboratories). In brief, the cells are incubated with the biotinylated mouse anti-BrdUrd antibody at 4 degrees C for 2 h after being exposed to denaturing solution and

10 then incubated with the streptavidin-peroxidase and diaminobenzidine. After counterstaining with hematoxylin, the cells are mounted for microscopic examination, and the BrdUrd-positive cells are counted. The BrdUrd index is calculated as a percent of the BrdUrd-positive cells to the total cell number. In addition, the simultaneous detection of the BrdUrd staining (nucleus) and the FITC uptake (cytoplasm) is performed for

15 individual cells by the concomitant use of bright field illumination and dark field-UV fluorescent illumination. See, Hayashida et al., J. Biol. Chem. 6:271(36):21985-21992 (1996).

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to

20 test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 39: Stimulation of Endothelial Migration**

This example will be used to explore the possibility that a polypeptide of the

25 invention may stimulate lymphatic endothelial cell migration.

Endothelial cell migration assays are performed using a 48 well microchemotaxis chamber (Neuroprobe Inc., Cabin John, MD; Falk, W., et al., J. Immunological Methods

1980;33:239-247). Polyvinylpyrrolidone-free polycarbonate filters with a pore size of 8 um (Nucleopore Corp. Cambridge, MA) are coated with 0.1% gelatin for at least 6 hours at room temperature and dried under sterile air. Test substances are diluted to appropriate concentrations in M199 supplemented with 0.25% bovine serum albumin (BSA), and 25  
5 ul of the final dilution is placed in the lower chamber of the modified Boyden apparatus. Subconfluent, early passage (2-6) HUVEC or BMEC cultures are washed and trypsinized for the minimum time required to achieve cell detachment. After placing the filter between lower and upper chamber,  $2.5 \times 10^5$  cells suspended in 50 ul M199 containing 1% FBS are seeded in the upper compartment. The apparatus is then incubated for 5 hours at  
10 37°C in a humidified chamber with 5% CO<sub>2</sub> to allow cell migration. After the incubation period, the filter is removed and the upper side of the filter with the non-migrated cells is scraped with a rubber policeman. The filters are fixed with methanol and stained with a Giemsa solution (Diff-Quick, Baxter, McGraw Park, IL). Migration is quantified by counting cells of three random high-power fields (40x) in each well, and all groups are  
15 performed in quadruplicate.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

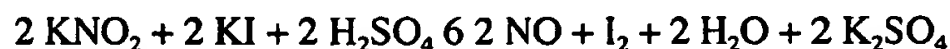
20

#### **Example 40: Stimulation of Nitric Oxide Production by Endothelial Cells**

Nitric oxide released by the vascular endothelium is believed to be a mediator of vascular endothelium relaxation. Thus, activity of a polypeptide of the invention can be assayed by determining nitric oxide production by endothelial cells in response to the  
25 polypeptide.

Nitric oxide is measured in 96-well plates of confluent microvascular endothelial cells after 24 hours starvation and a subsequent 4 hr exposure to various levels of a positive control (such as VEGF-1) and the polypeptide of the invention. Nitric oxide in the medium is determined by use of the Griess reagent to measure total nitrite after  
5 reduction of nitric oxide-derived nitrate by nitrate reductase. The effect of the polypeptide of the invention on nitric oxide release is examined on HUVEC.

Briefly, NO release from cultured HUVEC monolayer is measured with a NO-specific polarographic electrode connected to a NO meter (Iso-NO, World Precision Instruments Inc.) (1049). Calibration of the NO elements is performed according to the  
10 following equation:



The standard calibration curve is obtained by adding graded concentrations of  $\text{KNO}_2$  (0, 5, 10, 25, 50, 100, 250, and 500 nmol/L) into the calibration solution containing KI and  $\text{H}_2\text{SO}_4$ . The specificity of the Iso-NO electrode to NO is previously determined by  
15 measurement of NO from authentic NO gas (1050). The culture medium is removed and HUVECs are washed twice with Dulbecco's phosphate buffered saline. The cells are then bathed in 5 ml of filtered Krebs-Henseleit solution in 6-well plates, and the cell plates are kept on a slide warmer (Lab Line Instruments Inc.) To maintain the temperature at 37°C. The NO sensor probe is inserted vertically into the wells, keeping the tip of the electrode 2  
20 mm under the surface of the solution, before addition of the different conditions. S-nitroso acetyl penicillamin (SNAP) is used as a positive control. The amount of released NO is expressed as picomoles per  $1 \times 10^6$  endothelial cells. All values reported are means of four to six measurements in each group (number of cell culture wells). See, Leak *et al. Biochem. and Biophys. Res. Comm.* 217:96-105 (1995).

25 The studies described in this example tested activity of polypeptides of the invention. However, one skilled in the art could easily modify the exemplified studies to

test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 41: Effect of Polypeptides of the Invention on Cord Formation in**

**5    Angiogenesis**

Another step in angiogenesis is cord formation, marked by differentiation of endothelial cells. This bioassay measures the ability of microvascular endothelial cells to form capillary-like structures (hollow structures) when cultured *in vitro*.

CADMEC (microvascular endothelial cells) are purchased from Cell Applications, Inc. as proliferating (passage 2) cells and are cultured in Cell Applications' CADMEC Growth Medium and used at passage 5. For the *in vitro* angiogenesis assay, the wells of a 48-well cell culture plate are coated with Cell Applications' Attachment Factor Medium (200 ml/well) for 30 min. at 37°C. CADMEC are seeded onto the coated wells at 7,500 cells/well and cultured overnight in Growth Medium. The Growth Medium is then replaced with 300 mg Cell Applications' Chord Formation Medium containing control buffer or a polypeptide of the invention (0.1 to 100 ng/ml) and the cells are cultured for an additional 48 hr. The numbers and lengths of the capillary-like chords are quantitated through use of the Boeckeler VIA-170 video image analyzer. All assays are done in triplicate.

Commercial (R&D) VEGF (50 ng/ml) is used as a positive control. b-esteradiol (1 ng/ml) is used as a negative control. The appropriate buffer (without protein) is also utilized as a control.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.



**Example 42: Angiogenic Effect on Chick Chorioallantoic Membrane**

Chick chorioallantoic membrane (CAM) is a well-established system to examine angiogenesis. Blood vessel formation on CAM is easily visible and quantifiable. The ability of polypeptides of the invention to stimulate angiogenesis in CAM can be examined.

Fertilized eggs of the White Leghorn chick (*Gallus gallus*) and the Japanese quail (*Coturnix coturnix*) are incubated at 37.8°C and 80% humidity. Differentiated CAM of 16-day-old chick and 13-day-old quail embryos is studied with the following methods.

On Day 4 of development, a window is made into the egg shell of chick eggs. The embryos are checked for normal development and the eggs sealed with cellotape. They are further incubated until Day 13. Thermanox coverslips (Nunc, Naperville, IL) are cut into disks of about 5 mm in diameter. Sterile and salt-free growth factors are dissolved in distilled water and about 3.3 mg/ 5 ml are pipetted on the disks. After air-drying, the inverted disks are applied on CAM. After 3 days, the specimens are fixed in 3% glutaraldehyde and 2% formaldehyde and rinsed in 0.12 M sodium cacodylate buffer. They are photographed with a stereo microscope [Wild M8] and embedded for semi- and ultrathin sectioning as described above. Controls are performed with carrier disks alone.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 43: Angiogenesis Assay Using a Matrigel Implant in Mouse**

*In vivo* angiogenesis assay of a polypeptide of the invention measures the ability of an existing capillary network to form new vessels in an implanted capsule of murine

extracellular matrix material (Matrigel). The protein is mixed with the liquid Matrigel at 4 degree C and the mixture is then injected subcutaneously in mice where it solidifies. After 7 days, the solid "plug" of Matrigel is removed and examined for the presence of new blood vessels. Matrigel is purchased from Becton Dickinson Labware/Collaborative Biomedical Products.

When thawed at 4 degree C the Matrigel material is a liquid. The Matrigel is mixed with a polypeptide of the invention at 150 ng/ml at 4 degrees C and drawn into cold 3 ml syringes. Female C57Bl/6 mice approximately 8 weeks old are injected with the mixture of Matrigel and experimental protein at 2 sites at the midventral aspect of the abdomen (0.5 ml/site). After 7 days, the mice are sacrificed by cervical dislocation, the Matrigel plugs are removed and cleaned (i.e., all clinging membranes and fibrous tissue is removed). Replicate whole plugs are fixed in neutral buffered 10% formaldehyde, embedded in paraffin and used to produce sections for histological examination after staining with Masson's Trichrome. Cross sections from 3 different regions of each plug are processed. Selected sections are stained for the presence of vWF. The positive control for this assay is bovine basic FGF (150 ng/ml). Matrigel alone is used to determine basal levels of angiogenesis.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

#### **Example 44: Rescue of Ischemia in Rabbit Lower Limb Model**

To study the in vivo effects of polynucleotides and polypeptides of the invention on ischemia, a rabbit hindlimb ischemia model is created by surgical removal of one femoral arteries as described previously (Takeshita *et al.*, *Am J. Pathol* 147:1649-1660

(1995)). The excision of the femoral artery results in retrograde propagation of thrombus and occlusion of the external iliac artery. Consequently, blood flow to the ischemic limb is dependent upon collateral vessels originating from the internal iliac artery (Takeshita *et al. Am J. Pathol* 147:1649-1660 (1995)). An interval of 10 days is allowed for post-operative recovery of rabbits and development of endogenous collateral vessels. At 10 day post-operatively (day 0), after performing a baseline angiogram, the internal iliac artery of the ischemic limb is transfected with 500 mg naked expression plasmid containing a polynucleotide of the invention by arterial gene transfer technology using a hydrogel-coated balloon catheter as described (Riessen *et al. Hum Gene Ther.* 4:749-758 (1993); Leclerc *et al. J. Clin. Invest.* 90: 936-944 (1992)). When a polypeptide of the invention is used in the treatment, a single bolus of 500 mg polypeptide of the invention or control is delivered into the internal iliac artery of the ischemic limb over a period of 1 min. through an infusion catheter. On day 30, various parameters are measured in these rabbits: (a) BP ratio - The blood pressure ratio of systolic pressure of the ischemic limb to that of normal limb; (b) Blood Flow and Flow Reserve - Resting FL: the blood flow during undilated condition and Max FL: the blood flow during fully dilated condition (also an indirect measure of the blood vessel amount) and Flow Reserve is reflected by the ratio of max FL: resting FL; (c) Angiographic Score - This is measured by the angiogram of collateral vessels. A score is determined by the percentage of circles in an overlaying grid that with crossing opacified arteries divided by the total number in the rabbit thigh; (d) Capillary density - The number of collateral capillaries determined in light microscopic sections taken from hindlimbs.

The studies described in this example tested activity of polynucleotides and polypeptides of the invention. However, one skilled in the art could easily modify the exemplified studies to test the agonists, and/or antagonists of the invention.

**Example 45: Effect of Polypeptides of the Invention on Vasodilation**

Since dilation of vascular endothelium is important in reducing blood pressure, the ability of polypeptides of the invention to affect the blood pressure in spontaneously hypertensive rats (SHR) is examined. Increasing doses (0, 10, 30, 100, 300, and 900 mg/kg) of the polypeptides of the invention are administered to 13-14 week old spontaneously hypertensive rats (SHR). Data are expressed as the mean +/- SEM. Statistical analysis are performed with a paired t-test and statistical significance is defined as  $p < 0.05$  vs. the response to buffer alone.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 46: Rat Ischemic Skin Flap Model**

The evaluation parameters include skin blood flow, skin temperature, and factor VIII immunohistochemistry or endothelial alkaline phosphatase reaction. Expression of polypeptides of the invention, during the skin ischemia, is studied using in situ hybridization.

The study in this model is divided into three parts as follows:

- a) Ischemic skin
- b) Ischemic skin wounds
- c) Normal wounds

The experimental protocol includes:

- a) Raising a 3x4 cm, single pedicle full-thickness random skin flap (myocutaneous flap over the lower back of the animal).
- b) An excisional wounding (4-6 mm in diameter) in the ischemic skin (skin-flap).

c) Topical treatment with a polypeptide of the invention of the excisional wounds (day 0, 1, 2, 3, 4 post-wounding) at the following various dosage ranges: 1mg to 100 mg.

d) Harvesting the wound tissues at day 3, 5, 7, 10, 14 and 21 post-wounding for histological, immunohistochemical, and in situ studies.

5       The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

10    **Example 47: Peripheral Arterial Disease Model**

Angiogenic therapy using a polypeptide of the invention is a novel therapeutic strategy to obtain restoration of blood flow around the ischemia in case of peripheral arterial diseases. The experimental protocol includes:

15       a) One side of the femoral artery is ligated to create ischemic muscle of the hindlimb, the other side of hindlimb serves as a control.

b) a polypeptide of the invention, in a dosage range of 20 mg - 500 mg, is delivered intravenously and/or intramuscularly 3 times (perhaps more) per week for 2-3 weeks.

20       c) The ischemic muscle tissue is collected after ligation of the femoral artery at 1, 2, and 3 weeks for the analysis of expression of a polypeptide of the invention and histology. Biopsy is also performed on the other side of normal muscle of the contralateral hindlimb.

25       The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 48: Ischemic Myocardial Disease Model**

A polypeptide of the invention is evaluated as a potent mitogen capable of stimulating the development of collateral vessels, and restructuring new vessels after coronary artery occlusion. Alteration of expression of the polypeptide is investigated in situ. The experimental protocol includes:

- a) The heart is exposed through a left-side thoracotomy in the rat. Immediately, the left coronary artery is occluded with a thin suture (6-0) and the thorax is closed.
- b) a polypeptide of the invention, in a dosage range of 20 mg - 500 mg, is delivered intravenously and/or intramuscularly 3 times (perhaps more) per week for 2-4 weeks.

- c) Thirty days after the surgery, the heart is removed and cross-sectioned for morphometric and in situ analyzes.
- The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 49: Rat Corneal Wound Healing Model**

This animal model shows the effect of a polypeptide of the invention on neovascularization. The experimental protocol includes:

- a) Making a 1-1.5 mm long incision from the center of cornea into the stromal layer.
- b) Inserting a spatula below the lip of the incision facing the outer corner of the eye.

c) Making a pocket (its base is 1-1.5 mm from the edge of the eye).

d) Positioning a pellet, containing 50ng- 5ug of a polypeptide of the invention, within the pocket.

e) Treatment with a polypeptide of the invention can also be applied topically to the corneal wounds in a dosage range of 20mg - 500mg (daily treatment for five days).

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

10

#### **Example 50: Diabetic Mouse and Glucocorticoid-Impaired Wound Healing Models**

##### ***A. Diabetic db+/db+ Mouse Model.***

15 To demonstrate that a polypeptide of the invention accelerates the healing process, the genetically diabetic mouse model of wound healing is used. The full thickness wound healing model in the db+/db+ mouse is a well characterized, clinically relevant and reproducible model of impaired wound healing. Healing of the diabetic wound is dependent on formation of granulation tissue and re-epithelialization rather than contraction (Gartner, M.H. *et al.*, *J. Surg. Res.* 52:389 (1992); Greenhalgh, D.G. *et al.*, *Am. J. Pathol.* 136:1235 (1990)).

20 The diabetic animals have many of the characteristic features observed in Type II diabetes mellitus. Homozygous (db+/db+) mice are obese in comparison to their normal heterozygous (db+/-) littermates. Mutant diabetic (db+/db+) mice have a single autosomal recessive mutation on chromosome 4 (db+) (Coleman *et al.* *Proc. Natl. Acad. Sci. USA* 77:283-293 (1982)). Animals show polyphagia, polydipsia and polyuria. Mutant diabetic mice (db+/db+) have elevated blood glucose, increased or normal insulin

levels, and suppressed cell-mediated immunity (Mandel *et al.*, *J. Immunol.* 120:1375 (1978); Debray-Sachs, M. *et al.*, *Clin. Exp. Immunol.* 51(1):1-7 (1983); Leiter *et al.*, *Am. J. of Pathol.* 114:46-55 (1985)). Peripheral neuropathy, myocardial complications, and microvascular lesions, basement membrane thickening and glomerular filtration abnormalities have been described in these animals (Norido, F. *et al.*, *Exp. Neurol.* 83(2):221-232 (1984); Robertson *et al.*, *Diabetes* 29(1):60-67 (1980); Giacomelli *et al.*, *Lab Invest.* 40(4):460-473 (1979); Coleman, D.L., *Diabetes* 31 (Suppl):1-6 (1982)). These homozygous diabetic mice develop hyperglycemia that is resistant to insulin analogous to human type II diabetes (Mandel *et al.*, *J. Immunol.* 120:1375-1377 (1978)).

10       The characteristics observed in these animals suggests that healing in this model may be similar to the healing observed in human diabetes (Greenhalgh, *et al.*, *Am. J. of Pathol.* 136:1235-1246 (1990)).

Genetically diabetic female C57BL/KsJ (db+/db+) mice and their non-diabetic (db+/+m) heterozygous littermates are used in this study (Jackson Laboratories). The animals are purchased at 6 weeks of age and are 8 weeks old at the beginning of the study. Animals are individually housed and received food and water ad libitum. All manipulations are performed using aseptic techniques. The experiments are conducted according to the rules and guidelines of Human Genome Sciences, Inc. Institutional Animal Care and Use Committee and the Guidelines for the Care and Use of Laboratory Animals.

Wounding protocol is performed according to previously reported methods (Tsuboi, R. and Rifkin, D.B., *J. Exp. Med.* 172:245-251 (1990)). Briefly, on the day of wounding, animals are anesthetized with an intraperitoneal injection of Avertin (0.01 mg/mL), 2,2,2-tribromoethanol and 2-methyl-2-butanol dissolved in deionized water. The dorsal region of the animal is shaved and the skin washed with 70% ethanol solution and iodine. The surgical area is dried with sterile gauze prior to wounding. An 8 mm full-



thickness wound is then created using a Keyes tissue punch. Immediately following wounding, the surrounding skin is gently stretched to eliminate wound expansion. The wounds are left open for the duration of the experiment. Application of the treatment is given topically for 5 consecutive days commencing on the day of wounding. Prior to  
5 treatment, wounds are gently cleansed with sterile saline and gauze sponges.

Wounds are visually examined and photographed at a fixed distance at the day of surgery and at two day intervals thereafter. Wound closure is determined by daily measurement on days 1-5 and on day 8. Wounds are measured horizontally and vertically using a calibrated Jameson caliper. Wounds are considered healed if granulation tissue is  
10 no longer visible and the wound is covered by a continuous epithelium.

A polypeptide of the invention is administered using at a range different doses, from 4mg to 500mg per wound per day for 8 days in vehicle. Vehicle control groups received 50mL of vehicle solution.

Animals are euthanized on day 8 with an intraperitoneal injection of sodium  
15 pentobarbital (300mg/kg). The wounds and surrounding skin are then harvested for histology and immunohistochemistry. Tissue specimens are placed in 10% neutral buffered formalin in tissue cassettes between biopsy sponges for further processing.

Three groups of 10 animals each (5 diabetic and 5 non-diabetic controls) are evaluated: 1) Vehicle placebo control, 2) untreated group, and 3) treated group.

20 Wound closure is analyzed by measuring the area in the vertical and horizontal axis and obtaining the total square area of the wound. Contraction is then estimated by establishing the differences between the initial wound area (day 0) and that of post treatment (day 8). The wound area on day 1 is 64mm<sup>2</sup>, the corresponding size of the dermal punch. Calculations are made using the following formula:

25

$$[\text{Open area on day 8}] - [\text{Open area on day 1}] / [\text{Open area on day 1}]$$

Specimens are fixed in 10% buffered formalin and paraffin embedded blocks are sectioned perpendicular to the wound surface (5mm) and cut using a Reichert-Jung microtome. Routine hematoxylin-eosin (H&E) staining is performed on cross-sections of bisected wounds. Histologic examination of the wounds are used to assess whether the healing process and the morphologic appearance of the repaired skin is altered by treatment with a polypeptide of the invention. This assessment included verification of the presence of cell accumulation, inflammatory cells, capillaries, fibroblasts, re-epithelialization and epidermal maturity (Greenhalgh, D.G. *et al.*, *Am. J. Pathol.* 136:1235 (1990)). A calibrated lens micrometer is used by a blinded observer.

Tissue sections are also stained immunohistochemically with a polyclonal rabbit anti-human keratin antibody using ABC Elite detection system. Human skin is used as a positive tissue control while non-immune IgG is used as a negative control. Keratinocyte growth is determined by evaluating the extent of reepithelialization of the wound using a calibrated lens micrometer.

Proliferating cell nuclear antigen/cyclin (PCNA) in skin specimens is demonstrated by using anti-PCNA antibody (1:50) with an ABC Elite detection system. Human colon cancer can serve as a positive tissue control and human brain tissue can be used as a negative tissue control. Each specimen includes a section with omission of the primary antibody and substitution with non-immune mouse IgG. Ranking of these sections is based on the extent of proliferation on a scale of 0-8, the lower side of the scale reflecting slight proliferation to the higher side reflecting intense proliferation.

Experimental data are analyzed using an unpaired t test. A p value of < 0.05 is considered significant.

25

### ***B. Steroid Impaired Rat Model***

The inhibition of wound healing by steroids has been well documented in various *in vitro* and *in vivo* systems (Wahl, Glucocorticoids and Wound healing. In: Anti-Inflammatory Steroid Action: Basic and Clinical Aspects. 280-302 (1989); Wahlet *et al.*, *J. Immunol.* 115: 476-481 (1975); Werb *et al.*, *J. Exp. Med.* 147:1684-1694 (1978)).

5 Glucocorticoids retard wound healing by inhibiting angiogenesis, decreasing vascular permeability (Ebert *et al.*, *An. Intern. Med.* 37:701-705 (1952)), fibroblast proliferation, and collagen synthesis (Beck *et al.*, *Growth Factors.* 5: 295-304 (1991); Haynes *et al.*, *J. Clin. Invest.* 61: 703-797 (1978)) and producing a transient reduction of circulating monocytes (Haynes *et al.*, *J. Clin. Invest.* 61: 703-797 (1978); Wahl, "Glucocorticoids and

10 wound healing", In: Antiinflammatory Steroid Action: Basic and Clinical Aspects, Academic Press, New York, pp. 280-302 (1989)). The systemic administration of steroids to impaired wound healing is a well establish phenomenon in rats (Beck *et al.*, *Growth Factors.* 5: 295-304 (1991); Haynes *et al.*, *J. Clin. Invest.* 61: 703-797 (1978); Wahl, "Glucocorticoids and wound healing", In: Antiinflammatory Steroid Action: Basic and

15 Clinical Aspects, Academic Press, New York, pp. 280-302 (1989); Pierce *et al.*, *Proc. Natl. Acad. Sci. USA* 86: 2229-2233 (1989)).

To demonstrate that a polypeptide of the invention can accelerate the healing process, the effects of multiple topical applications of the polypeptide on full thickness excisional skin wounds in rats in which healing has been impaired by the systemic

20 administration of methylprednisolone is assessed.

Young adult male Sprague Dawley rats weighing 250-300 g (Charles River Laboratories) are used in this example. The animals are purchased at 8 weeks of age and are 9 weeks old at the beginning of the study. The healing response of rats is impaired by the systemic administration of methylprednisolone (17mg/kg/rat intramuscularly) at the

25 time of wounding. Animals are individually housed and received food and water *ad libitum*. All manipulations are performed using aseptic techniques. This study is

conducted according to the rules and guidelines of Human Genome Sciences, Inc. Institutional Animal Care and Use Committee and the Guidelines for the Care and Use of Laboratory Animals.

The wounding protocol is followed according to section A, above. On the day of  
5 wounding, animals are anesthetized with an intramuscular injection of ketamine (50 mg/kg) and xylazine (5 mg/kg). The dorsal region of the animal is shaved and the skin washed with 70% ethanol and iodine solutions. The surgical area is dried with sterile gauze prior to wounding. An 8 mm full-thickness wound is created using a Keyes tissue punch. The wounds are left open for the duration of the experiment. Applications of the  
10 testing materials are given topically once a day for 7 consecutive days commencing on the day of wounding and subsequent to methylprednisolone administration. Prior to treatment, wounds are gently cleansed with sterile saline and gauze sponges.

Wounds are visually examined and photographed at a fixed distance at the day of wounding and at the end of treatment. Wound closure is determined by daily measurement  
15 on days 1-5 and on day 8. Wounds are measured horizontally and vertically using a calibrated Jameson caliper. Wounds are considered healed if granulation tissue is no longer visible and the wound is covered by a continuous epithelium.

The polypeptide of the invention is administered using at a range different doses, from 4mg to 500mg per wound per day for 8 days in vehicle. Vehicle control groups  
20 received 50mL of vehicle solution.

Animals are euthanized on day 8 with an intraperitoneal injection of sodium pentobarbital (300mg/kg). The wounds and surrounding skin are then harvested for histology. Tissue specimens are placed in 10% neutral buffered formalin in tissue cassettes between biopsy sponges for further processing.

Four groups of 10 animals each (5 with methylprednisolone and 5 without glucocorticoid) are evaluated: 1) Untreated group 2) Vehicle placebo control 3) treated groups.

Wound closure is analyzed by measuring the area in the vertical and horizontal axis and obtaining the total area of the wound. Closure is then estimated by establishing the differences between the initial wound area (day 0) and that of post treatment (day 8). The wound area on day 1 is 64mm<sup>2</sup>, the corresponding size of the dermal punch. Calculations are made using the following formula:

$$[\text{Open area on day 8}] - [\text{Open area on day 1}] / [\text{Open area on day 1}]$$

Specimens are fixed in 10% buffered formalin and paraffin embedded blocks are sectioned perpendicular to the wound surface (5mm) and cut using an Olympus microtome. Routine hematoxylin-eosin (H&E) staining is performed on cross-sections of bisected wounds. Histologic examination of the wounds allows assessment of whether the healing process and the morphologic appearance of the repaired skin is improved by treatment with a polypeptide of the invention. A calibrated lens micrometer is used by a blinded observer to determine the distance of the wound gap.

Experimental data are analyzed using an unpaired t test. A p value of < 0.05 is considered significant.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

**Example 51: Lymphadema Animal Model**

The purpose of this experimental approach is to create an appropriate and consistent lymphedema model for testing the therapeutic effects of a polypeptide of the invention in lymphangiogenesis and re-establishment of the lymphatic circulatory system  
5 in the rat hind limb. Effectiveness is measured by swelling volume of the affected limb, quantification of the amount of lymphatic vasculature, total blood plasma protein, and histopathology. Acute lymphedema is observed for 7-10 days. Perhaps more importantly, the chronic progress of the edema is followed for up to 3-4 weeks.

Prior to beginning surgery, blood sample is drawn for protein concentration  
10 analysis. Male rats weighing approximately ~350g are dosed with Pentobarbital. Subsequently, the right legs are shaved from knee to hip. The shaved area is swabbed with gauze soaked in 70% EtOH. Blood is drawn for serum total protein testing. Circumference and volumetric measurements are made prior to injecting dye into paws after marking 2 measurement levels (0.5 cm above heel, at mid-pt of dorsal paw). The  
15 intradermal dorsum of both right and left paws are injected with 0.05 ml of 1% Evan's Blue. Circumference and volumetric measurements are then made following injection of dye into paws.

Using the knee joint as a landmark, a mid-leg inguinal incision is made circumferentially allowing the femoral vessels to be located. Forceps and hemostats are  
20 used to dissect and separate the skin flaps. After locating the femoral vessels, the lymphatic vessel that runs along side and underneath the vessel(s) is located. The main lymphatic vessels in this area are then electrically coagulated suture ligated.

Using a microscope, muscles in back of the leg (near the semitendinosus and adductors) are bluntly dissected. The popliteal lymph node is then located. The 2  
25 proximal and 2 distal lymphatic vessels and distal blood supply of the popliteal node are

then and ligated by suturing. The popliteal lymph node, and any accompanying adipose tissue, is then removed by cutting connective tissues.

Care is taken to control any mild bleeding resulting from this procedure. After lymphatics are occluded, the skin flaps are sealed by using liquid skin (Vetbond) (AJ  
5 Buck). The separated skin edges are sealed to the underlying muscle tissue while leaving a gap of ~0.5 cm around the leg. Skin also may be anchored by suturing to underlying muscle when necessary.

To avoid infection, animals are housed individually with mesh (no bedding). Recovering animals are checked daily through the optimal edematous peak, which  
10 typically occurred by day 5-7. The plateau edematous peak are then observed. To evaluate the intensity of the lymphedema, the circumference and volumes of 2 designated places on each paw before operation and daily for 7 days are measured. The effect plasma proteins on lymphedema is determined and whether protein analysis is a useful testing  
15 perimeter is also investigated. The weights of both control and edematous limbs are evaluated at 2 places. Analysis is performed in a blind manner.

Circumference Measurements: Under brief gas anesthetic to prevent limb movement, a cloth tape is used to measure limb circumference. Measurements are done at the ankle bone and dorsal paw by 2 different people then those 2 readings are averaged. Readings are taken from both control and edematous limbs.

20 Volumetric Measurements: On the day of surgery, animals are anesthetized with Pentobarbital and are tested prior to surgery. For daily volumetrics animals are under brief halothane anesthetic (rapid immobilization and quick recovery), both legs are shaved and equally marked using waterproof marker on legs. Legs are first dipped in water, then dipped into instrument to each marked level then measured by Buxco edema  
25 software(Chen/Victor). Data is recorded by one person, while the other is dipping the limb to marked area.

Blood-plasma protein measurements: Blood is drawn, spun, and serum separated prior to surgery and then at conclusion for total protein and  $\text{Ca}^{2+}$  comparison.

Limb Weight Comparison: After drawing blood, the animal is prepared for tissue collection. The limbs are amputated using a quillitine, then both experimental and control  
5 legs are cut at the ligature and weighed. A second weighing is done as the tibio-cacaneal joint is disarticulated and the foot is weighed.

Histological Preparations: The transverse muscle located behind the knee (popliteal) area is dissected and arranged in a metal mold, filled with freezeGel, dipped into cold methylbutane, placed into labeled sample bags at  $-80^{\circ}\text{C}$  until sectioning. Upon  
10 sectioning, the muscle is observed under fluorescent microscopy for lymphatics..

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

15

**Example 52: Suppression of TNF alpha-induced adhesion molecule expression by a Polypeptide of the Invention**

The recruitment of lymphocytes to areas of inflammation and angiogenesis involves specific receptor-ligand interactions between cell surface adhesion molecules  
20 (CAMs) on lymphocytes and the vascular endothelium. The adhesion process, in both normal and pathological settings, follows a multi-step cascade that involves intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1 (VCAM-1), and endothelial leukocyte adhesion molecule-1 (E-selectin) expression on endothelial cells (EC). The expression of these molecules and others on the vascular endothelium  
25 determines the efficiency with which leukocytes may adhere to the local vasculature and extravasate into the local tissue during the development of an inflammatory response. The



local concentration of cytokines and growth factor participate in the modulation of the expression of these CAMs.

Tumor necrosis factor alpha (TNF- $\alpha$ ), a potent proinflammatory cytokine, is a stimulator of all three CAMs on endothelial cells and may be involved in a wide variety of inflammatory responses, often resulting in a pathological outcome.

The potential of a polypeptide of the invention to mediate a suppression of TNF- $\alpha$  induced CAM expression can be examined. A modified ELISA assay which uses ECs as a solid phase absorbent is employed to measure the amount of CAM expression on TNF- $\alpha$  treated ECs when co-stimulated with a member of the FGF family of proteins.

To perform the experiment, human umbilical vein endothelial cell (HUVEC) cultures are obtained from pooled cord harvests and maintained in growth medium (EGM-2; Clonetics, San Diego, CA) supplemented with 10% FCS and 1% penicillin/streptomycin in a 37 degree C humidified incubator containing 5% CO<sub>2</sub>. HUVECs are seeded in 96-well plates at concentrations of  $1 \times 10^4$  cells/well in EGM medium at 37 degree C for 18-24 hrs or until confluent. The monolayers are subsequently washed 3 times with a serum-free solution of RPMI-1640 supplemented with 100 U/ml penicillin and 100 mg/ml streptomycin, and treated with a given cytokine and/or growth factor(s) for 24 h at 37 degree C. Following incubation, the cells are then evaluated for CAM expression.

Human Umbilical Vein Endothelial cells (HUVECs) are grown in a standard 96 well plate to confluence. Growth medium is removed from the cells and replaced with 90  $\mu$ l of 199 Medium (10% FBS). Samples for testing and positive or negative controls are added to the plate in triplicate (in 10  $\mu$ l volumes). Plates are incubated at 37 degree C for either 5 h (selectin and integrin expression) or 24 h (integrin expression only). Plates are aspirated to remove medium and 100  $\mu$ l of 0.1% paraformaldehyde-PBS(with Ca<sup>++</sup> and Mg<sup>++</sup>) is added to each well. Plates are held at 4°C for 30 min.

Fixative is then removed from the wells and wells are washed 1X with PBS(+Ca,Mg)+0.5% BSA and drained. Do not allow the wells to dry. Add 10  $\mu$ l of diluted primary antibody to the test and control wells. Anti-ICAM-1-Biotin, Anti-VCAM-1-Biotin and Anti-E-selectin-Biotin are used at a concentration of 10  $\mu$ g/ml (1:10 dilution of 0.1 mg/ml stock antibody). Cells are incubated at 37°C for 30 min. in a humidified environment. Wells are washed X3 with PBS(+Ca,Mg)+0.5% BSA.

Then add 20  $\mu$ l of diluted ExtrAvidin-Alkaline Phosphatase (1:5,000 dilution) to each well and incubated at 37°C for 30 min. Wells are washed X3 with PBS(+Ca,Mg)+0.5% BSA. 1 tablet of p-Nitrophenol Phosphate pNPP is dissolved in 5 ml of glycine buffer (pH 10.4). 100  $\mu$ l of pNPP substrate in glycine buffer is added to each test well. Standard wells in triplicate are prepared from the working dilution of the ExtrAvidin-Alkaline Phosphatase in glycine buffer: 1:5,000 ( $10^0$ ) >  $10^{-0.5}$  >  $10^{-1}$  >  $10^{-1.5}$ . 5  $\mu$ l of each dilution is added to triplicate wells and the resulting AP content in each well is 5.50 ng, 1.74 ng, 0.55 ng, 0.18 ng. 100  $\mu$ l of pNPP reagent must then be added to each of the standard wells. The plate must be incubated at 37°C for 4h. A volume of 50  $\mu$ l of 3M NaOH is added to all wells. The results are quantified on a plate reader at 405 nm. The background subtraction option is used on blank wells filled with glycine buffer only. The template is set up to indicate the concentration of AP-conjugate in each standard well [5.50 ng; 1.74 ng; 0.55 ng; 0.18 ng]. Results are indicated as amount of bound AP-conjugate in each sample.

The studies described in this example tested activity of a polypeptide of the invention. However, one skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides (e.g., gene therapy), agonists, and/or antagonists of the invention.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous

modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other  
5 disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorporated herein by reference in their entireties.

## INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 13bis)

<b>A.</b> The indications made below relate to the microorganism referred to in the description on page <u>89</u> , line <u>N/A</u>	
<b>B. IDENTIFICATION OF DEPOSIT</b> Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depositary institution <u>American Type Culture Collection</u>	
Address of depositary institution <i>(including postal code and country)</i> <u>10801 University Boulevard</u> <u>Manassas, Virginia 20110-2209</u> <u>United States of America</u>	
Date of deposit <u>26 April 1999</u>	Accession Number <u>203959</u>
<b>C. ADDITIONAL INDICATIONS</b> <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
<b>D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE</b> <i>(if the indications are not for all designated States)</i>	
<u>Europe</u>  In respect to those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28 (4) EPC).	
<b>E. SEPARATE FURNISHING OF INDICATIONS</b> <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the International Bureau later <i>(specify the general nature of the indications e.g., "Accession Number of Deposit")</i>	

<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For receiving Office use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received with the international application</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>	<div style="text-align: right; font-weight: bold; margin-bottom: 5px;">For International Bureau use only</div> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"><input type="checkbox"/> This sheet was received by the International Bureau on:</div> <div style="border: 1px solid black; padding: 5px; min-height: 40px;">Authorized officer</div>
---	--

**ATCC Deposit No. 203959****CANADA**

The applicant requests that, until either a Canadian patent has been issued on the basis of an application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the Commissioner of Patents only authorizes the furnishing of a sample of the deposited biological material referred to in the application to an independent expert nominated by the Commissioner, the applicant must, by a written statement, inform the International Bureau accordingly before completion of technical preparations for publication of the international application.

**NORWAY**

The applicant hereby requests that the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on the list of recognized experts drawn up by the Norwegian Patent Office or any person approved by the applicant in the individual case.

**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

**FINLAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Regulations), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

**UNITED KINGDOM**

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for the international publication of the application.

**ATCC Deposit No.: 203959**

## **DENMARK**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent Office or any person by the applicant in the individual case.

## **SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent Office or any person approved by a applicant in the individual case.

## **NETHERLANDS**

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in the 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

***What Is Claimed Is:***

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 95% identical to a sequence selected from the group
- 5 consisting of:
- (a) a polynucleotide fragment of SEQ ID NO:X or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:Y or a
  - 10 polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:Y or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - 15 (d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:Y or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X;
  - (e) a polynucleotide encoding a polypeptide of SEQ ID NO:Y or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X,
  - 20 having biological activity;
  - (f) a polynucleotide which is a variant of SEQ ID NO:X;
  - (g) a polynucleotide which is an allelic variant of SEQ ID NO:X;
  - (h) a polynucleotide which encodes a species homologue of the SEQ ID NO:Y;
  - 25 (i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not

hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the  
5 polynucleotide fragment comprises a nucleotide sequence encoding a secreted protein.

3. The isolated nucleic acid molecule of claim 1, wherein the  
polynucleotide fragment comprises a nucleotide sequence encoding the sequence  
10 identified as SEQ ID NO:Y or the polypeptide encoded by the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

4. The isolated nucleic acid molecule of claim 1, wherein the  
polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:X  
15 or the cDNA sequence included in ATCC Deposit No:Z, which is hybridizable to SEQ ID NO:X.

5. The isolated nucleic acid molecule of claim 2, wherein the nucleotide  
sequence comprises sequential nucleotide deletions from either the C-terminus or the  
20 N-terminus.

6. The isolated nucleic acid molecule of claim 3, wherein the nucleotide  
sequence comprises sequential nucleotide deletions from either the C-terminus or the  
N-terminus.



7. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.
8. A method of making a recombinant host cell comprising the isolated  
5 nucleic acid molecule of claim 1.
9. A recombinant host cell produced by the method of claim 8.
10. The recombinant host cell of claim 9 comprising vector sequences.  
10
11. An isolated polypeptide comprising an amino acid sequence at least 95% identical to a sequence selected from the group consisting of:
- (a) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - 15 (b) a polypeptide fragment of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z, having biological activity;
  - (c) a polypeptide domain of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - (d) a polypeptide epitope of SEQ ID NO:Y or the encoded sequence included  
20 in ATCC Deposit No:Z;
  - (e) a secreted form of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - (f) a full length protein of SEQ ID NO:Y or the encoded sequence included in ATCC Deposit No:Z;
  - 25 (g) a variant of SEQ ID NO:Y;
  - (h) an allelic variant of SEQ ID NO:Y; or

(i) a species homologue of the SEQ ID NO:Y.

12. The isolated polypeptide of claim 11, wherein the secreted form or the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

5

13. An isolated antibody that binds specifically to the isolated polypeptide of claim 11.

14. A recombinant host cell that expresses the isolated polypeptide of  
10 claim 11.

15. A method of making an isolated polypeptide comprising:

(a) culturing the recombinant host cell of claim 14 under conditions such that said polypeptide is expressed; and

15 (b) recovering said polypeptide.

16. The polypeptide produced by claim 15.

17. A method for preventing, treating, or ameliorating a medical condition,  
20 comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 11 or the polynucleotide of claim 1.

18. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

25 (a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

19. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 11 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

10

20. A method for identifying a binding partner to the polypeptide of claim 11 comprising:

(a) contacting the polypeptide of claim 11 with a binding partner; and

(b) determining whether the binding partner effects an activity of the polypeptide.

15

21. The gene corresponding to the cDNA sequence of SEQ ID NO:Y.

22. A method of identifying an activity in a biological assay, wherein the method comprises:

20

(a) expressing SEQ ID NO:X in a cell;

(b) isolating the supernatant;

(c) detecting an activity in a biological assay; and

(d) identifying the protein in the supernatant having the activity.

25

23. The product produced by the method of claim 20.

<110> Human Genome Sciences, Inc.

<120> 50 Human Secreted Proteins

<130> PS543PCT

<140> Unassigned

<141> 2000-04-06

<150> 60/128,703

<151> 1999-04-09

<150> 60/176,068

<151> 2000-01-20

<160> 152

<170> PatentIn Ver. 2.0

<210> 1

<211> 733

<212> DNA

<213> Homo sapiens

<400> 1

gggatccgga	gcccaaattct	tctgacaaaa	ctcacacatg	cccaccgtgc	ccagcacctg	60
aattcgaggg	tgcaccgtca	gtcttcctct	tccccccaaa	acccaaggac	accctcatga	120
tctcccggac	tcctgaggtc	acatgcgtgg	tggaggacgt	aagccacgaa	gaccctgagg	180
tcaagttcaa	ctggtacgtg	gacggcgtgg	aggtgcataa	tgccaagaca	aagccgcggg	240
aggagcagta	caacagcacg	taccgtgtgg	tcagcgtcct	caccgtcctg	caccaggact	300
ggctgaatgg	caaggagtac	aagtgcgaag	tctccaacaa	agccctccca	acccccatcg	360
agaaaaccat	ctccaaagcc	aaagggcagc	cccgagaacc	acaggtgtac	accctgcccc	420
catcccggga	tgagctgacc	aagaaccagg	tcagcctgac	ctgcctgggc	aaaggcttct	480
atccaagcga	catcgccgtg	gagtgggaga	gcaatgggca	gccggagAAC	aactacaaga	540
ccacgcctcc	cgtgctggac	tccgacggct	ccttcttcct	ctacagcaag	ctcaccgtgg	600
acaagagcag	gtggcagcag	gggaacgtct	tctcatgctc	cgtgatgcat	gaggctctgc	660
acaaccacta	cacgcagaag	agcctctccc	tgtctccggg	taaatgagtg	cgacggccgc	720
gactctagag	gat					733

<210> 2

<211> 5

<212> PRT

<213> Homo sapiens

<220>

<221> Site

<222> (3)

<223> Xaa equals any of the twenty naturally occurring L-amino acids

<400> 2

Trp Ser Xaa Trp Ser

1

5

<210> 3  
<211> 86  
<212> DNA  
<213> Homo sapiens

<400> 3  
gcgcctcgag atttccccga aatctagatt tccccgaaat gatttccccg aaatgatttc 60  
cccgaaatat ctgccatctc aattag 86

<210> 4  
<211> 27  
<212> DNA  
<213> Homo sapiens

<400> 4  
gcggcaagct ttttgcaaag cctaggc 27

<210> 5  
<211> 271  
<212> DNA  
<213> Homo sapiens

<400> 5  
ctcgagattt ccccgaaatc tagatttccc cgaaatgatt tccccgaaat gatttccccg 60  
aaatatctgc catctcaatt agtcagcaac catagtcccc cccctaactc cgcccatccc 120  
gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa ttttttttat 180  
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 240  
ttttggaggc ctaggctttt gcaaaaagct t 271

<210> 6  
<211> 32  
<212> DNA  
<213> Homo sapiens

<400> 6  
gcgctcgagg gatgacagcg atagaacccc gg 32

<210> 7  
<211> 31  
<212> DNA  
<213> Homo sapiens

<400> 7  
gcgaagcttc gcgactcccc ggatccgcct c 31

<210> 8  
<211> 12  
<212> DNA  
<213> Homo sapiens

<400> 8  
ggggactttc cc

12

<210> 9  
<211> 73  
<212> DNA  
<213> Homo sapiens

<400> 9  
gcggcctcga ggggactttc ccggggactt tccggggact ttccgggact ttccatcctg  
ccatctcaat tag

60

73

<210> 10  
<211> 256  
<212> DNA  
<213> Homo sapiens

<400> 10  
ctcgagggga ctttcccggg gactttccgg ggactttccg ggactttcca tctgccatct  
caattagtca gcaaccatag tcccgccct aactccgcc atcccgcccc taactccgcc  
cagttccgcc cattctccgc cccatggctg actaattttt tttatttatg cagaggccga  
ggccgcctcg gcctctgagc tattccagaa gtagtgagga ggcttttttg gaggcctagg  
cttttgcaaa aagctt

60

120

180

240

256

<210> 11  
<211> 1808  
<212> DNA  
<213> Homo sapiens

<400> 11  
ggcacgaggg cccaactggg gtatctgccc agcccttccc tccagtgttc ctccctagca  
ttcaagtgt acccagcaga aatataatgc cagttctgca tctaatttta aatttcctaa  
tagccatgct aaaaaagtaa aaagaaacag ttgaaattaa ttttaataat gtatctcgcc  
caatataatc aaactgttat ttcaacatgc aatttttgaa aaatgtgaga tgttttacat  
tctctttttc acacttagtc ttcaatatct ggtgtgtatt ttacacttct agcacctctc  
aattcacacc agccacattt aaaatgctca gtagccgcat gcagcttgtg cagccttact  
taggtccagt gtcagcctgg ggccgtgctg gggagctgag ctgtgtccat gggggccctg  
gtgagcctcc agtgctgtgc tgagggtgag ggcaccagc gcaagcccca gcgcagctgc  
ctgtccagcc ctgcctcagc tccttgacaa gaatgagctt gaagtgtctc cctctgctcg  
gttctagtgc tgcctctgac agttttgctc tttggttcac aaggcttcct aaactcagga  
ttgctttgtg ggtctgttct ttgggatgag aggccagggc aggtgcaagc tatgccttca  
ctcctttgtc cctgccttcc tccttcctgt accccagcta cagccccgtt gggaagaatg  
tgctgtagt cacaagcat tgcagatgtg tctaagaaat gccacgctgg ggccgggctg  
ggtgctcaca cctgtaatcc cagcacttcg agaggctgag gcaggcggat tgcctgagct  
caggatttca agaccaccct gggtaacatg gtgaaacct gtcttacta aaaatacaaa  
aaaattagct ggggtgtgtg gtacacgcct atagtcaccag ctacatggga ggctgagcag  
gagaatcgct tgaaccggga ggtggaggtt gcagtaagcc gagattgcac cactgcactc  
cagcctgggc gacaagagcg aaactccatc tcaaaaaaaa aaaaaaaac aaaaaaaac  
tattctgtat aggtaatgaa cgattttcaa aaatcttaga tgaatcaatg agataacagc  
caaacactca ctcaactggc atctactcaa aaagctattg gtttctcatg attagctttg  
ccaaaccatc tccctagatt atttttcttt ttaaatttca ctttgcatct ggttaatcat

60

120

180

240

300

360

420

480

540

600

660

720

780

840

900

960

1020

1080

1140

1200

1260

tccctgggaa	agcacacggg	gcaggtgggc	ctccttgtct	tcactttgcc	attccctatc	1320
tgatgaattc	tgaacctcag	tttttcatcc	aagaactgga	gttaaaacac	ctgcactatt	1380
atacagggcg	tgaggctggt	gtcatgataa	tcaatgagct	gatgtgtggt	tgaagctctt	1440
atctgactcc	atagatagtt	ttaaactacc	taagtataaa	ttcagcagct	ttgcttaaga	1500
tttaaagcag	gtattataaa	tatgcattcc	tttgccaatc	ttttaataga	aggacaggcc	1560
tattcttttg	aagatggatc	tgctgatgag	agctcccctt	tgtctacttt	acatcaacca	1620
cacccttatt	tcattgtttt	gtgattccag	tgttgggttc	tttaaagtaa	aggaagaatt	1680
tagatatattg	ccgagccatt	ctgaatatag	aaacttccta	gatcgcatat	cccttgatct	1740
tttatcgтта	attagatgag	agtaaattctc	gtgccgaatt	cgatatcaag	cttatcgata	1800
ccgtcgac						1808

&lt;210&gt; 12

&lt;211&gt; 1743

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 12

ggagaagaga	gaggcaaggg	caggagtga	ggagagagct	gaagcctggg	gctccgagat	60
ggtcagagga	tgggagacgg	ggcagtga	caaggcttct	tgtatcttca	gcagcagcag	120
acgtttggaa	agaaatggcg	ccgcttcggc	gcctcactgt	atggagggtc	ggactgcgcc	180
ttggcccggc	tggagctgca	ggagggcccg	gagaagcctc	gtcgggtgtga	ggctgcccgg	240
aaggtcatcc	gcctcagtga	ctgcctgcgg	gtggccgagg	ccggcggaga	ggccagcagc	300
ccccgggaca	ccagtgcctt	cttcctggag	accaaggagc	gcctgtacct	cctggcggcc	360
cctgcagcgg	agcgcggcga	ctgggtgcag	gccatctgcc	tcctggcctt	ccccgggcag	420
aggaaggagc	tctcggggcc	agagggaaa	cagagccggc	ccttcatggg	gaaatcta	480
tgtacagctt	gcgcagtcac	agtcggcccc	cacaaggaat	ttgctgtgac	catgagacct	540
acagaagcca	gtgaaaggtg	ccacctgcgg	gggtcctata	ccctccgggc	tggggaaagt	600
gccctggagc	tgtgggggtg	gcccgaagcca	gggaccagc	tgtacgactg	gccctacagg	660
tttctgcggc	gctttggggc	ggacaaggcc	gtcgtgcgt	ctctggagag	ggcaactttg	720
agttcgaaac	ccggcaaggc	aatgagatct	tcttggccct	ggaagaggcc	atctctgccc	780
agaagaatgc	tgcacccgct	acaccccac	cgcagccagc	cacaatcccc	gcctcgctgc	840
ccgggcctga	tagccctac	tctcggccgc	atgactcact	gccggccgct	tcacccacca	900
caccggtgcc	tgctccacgg	cctcggggcc	aggaggggga	gtatgccgtg	cccttcgatg	960
cggtggccc	ttccttgggg	aagaacttca	ggggcatctt	ggcagtccct	cctcagctcc	1020
tggccgaccc	ttctgtacga	cagcattgag	gagaccctgc	cccctcgacc	tgaccacata	1080
tacgatgagc	ccgagggagt	ggctgccctg	tccctctatg	acagcccgca	ggagccccgg	1140
ggtgaggcat	ggaggaggca	ggcgacagct	gacagggacc	ctgctggcct	ccagcatgtc	1200
cagccagctg	ggcaggattt	ctctgcttct	ggctggcagc	caggaactga	gtatgacaat	1260
gttgactaa	agaaaggccc	aaagtgacag	aggcagcaga	gggatggtcc	accgcccctt	1320
ggcttctgct	ggtgactcct	cctggccact	gcatacaga	aacctcctct	gccccttctg	1380
gagcccaggg	cctggcctgt	cttcgttggg	gctgataaat	tgcctctccc	agggcctgct	1440
gggtgagtca	ccatcccaaa	gcaggaaggg	tgccctggag	agaaccaccc	tcctcctact	1500
ctttttccac	ttcctcctct	ttctttcccc	agctgaggag	gaacctgggg	catttagggc	1560
agaggacaaa	aggatgtcag	caattgcttg	ggctgcttgg	ctatgcaagc	ctcctgcctg	1620
ctgatggcca	cttcagggac	agcctggggc	caggcaccca	gggggatggc	ggcagcttcc	1680
tgcacctttc	agattttcttg	gtggcaatta	aagcattttc	agaacaaaaa	aaaaaaaaaa	1740
aaa						1743

&lt;210&gt; 13

&lt;211&gt; 627

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 13

gaattcggca	cgaggaagtt	caacctggag	agatggaaaa	tcagctctca	taactaagtt	60
aatttagtat	aaaaatagaa	ttgatagtga	gggtataaag	tgtaaccatc	agttaaacct	120
ctcctgtcat	tcctggcttc	cttgcttcag	aattgaaatg	gaagtggggg	tgccctact	180
ctgtagaatc	tgggactggg	caaagtgttg	tgtggcctcc	ttaaactagc	tgttatgtta	240
tgattttatt	ctttgtgagt	taattagaat	aaagtcattt	tcttccaagg	tatggttcat	300
ttagtctata	gtctcctggg	tatgaaatta	gcacccctcc	agatctgaca	gctccctgag	360
gggttatata	aggagtagct	cacttgcaaa	cataacatga	tgctcaattc	aaagggttcta	420
gctgctgtgg	gtctggagtt	atgtaagctg	agcagtagca	ggaaactctg	cttaaaatat	480
tgtttaaagc	tgtgctgtgct	ggcatgtgcc	tgtaatccca	gctacttggg	agactgaggc	540
gggagaatca	cttgaaccca	ggaggtggag	gttgcagtga	actgagacca	tgccactgca	600
ctcactgaca	gagcaagact	ctgtctc				627

&lt;210&gt; 14

&lt;211&gt; 535

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (11)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (13)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (162)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (422)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 14

tgaaatatat	ntncatatgg	araaggggtct	cccttggtcg	ctggatatgg	tttggctgtg	60
tccctacca	aatcttatct	tgaattgtat	ttcccataat	tcttaggtgt	aatgggaggg	120
acctggtkgg	gaggtaattg	aatcatggga	gccaggtctt	tncctgtgct	gttctgggtg	180
tggtgaataa	gtctcatgag	atctgatggg	tttataaagg	ggagttcccc	gatcatgctc	240
tcttgccctgc	tgccatgtaa	gacatgcctt	tgcttctcct	ttgccttcca	ccatgattgt	300
catgtcttcc	tagccaagtg	gaactgtgag	tccattaaac	ctctttttct	ttataaatta	360
cgtgggtcca	aggccaggtg	cagtggctca	cgcctgtaat	cccaacactt	tgggagtctg	420
anggcaggcg	gatcacgagg	tcaggagttc	aagaccagcc	tgaaccaaca	tgatgaaacc	480
ctgtctctac	taaaaataca	aaaatttagc	caggcgtggg	ggcacacacc	tgtaa	535

&lt;210&gt; 15

&lt;211&gt; 719

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens



## &lt;400&gt; 15

ccacgcgtcc	gcctgcacgc	aacagaagaa	gtcctcctgt	caggcgccag	agaggaagaa	60
gggatcgtct	gtctcgacat	aattccatta	gtcaagatga	aaactatcac	catctccctt	120
acgcacagca	gcaagcaata	gggagcctcg	agccttccac	cctccgaatg	tatctccccg	180
tctgctacat	cctgctgctc	atccacccca	gcagaatgca	gtcatggttg	acatacatga	240
tcaggtacaa	taacagaatg	tcttggaat	acagaagctg	aaattttgtg	gcagtttgag	300
atTTTTTTTT	tggctTTTTg	ttgcaaattt	cttttagttt	ttacctaaaa	tgaagttatt	360
taagttcctc	attcctttta	agaattgaga	tttaaattct	acacagcttc	catagaagaa	420
agttgaactt	aaaaaattag	tatcggccca	gcacgggtgc	tcatgcctgt	aatcccagca	480
ctttgggagg	ctgaggcagg	cacatcactt	gagcttaggc	attcgagaca	agcctagcac	540
agtgaacccc	cgtctctact	ggaagtacga	aaattagctg	gatgtgggtg	tacatgttta	600
tgatcccagt	tgatgcagaa	gaatcacttg	aactcaggag	gtggagggtg	cagtgaagtca	660
agatcgcacc	actgcactgg	gtggcagaac	gagacttcgt	ctcaaaaaaa	aaaaaaaaaa	719

## &lt;210&gt; 16

&lt;211&gt; 901

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

## &lt;400&gt; 16

gctcgtgccg	cacgggagge	agaggttgca	atgagtcgag	atcgtgccat	tgcactccag	60
cctgggcagc	agagcgagac	tccggctcaa	aataataata	ataaattaga	gatgggggtct	120
cactatTTTT	cccaggctgg	TTTTgaactc	ctgagctcaa	gggatcctct	cgcctcagcc	180
tcccaaagca	ctgggatcac	aggtgtgagg	tgccatgcct	ggcccacacc	agctgtgttt	240
aatcaatgct	gggcagccct	gcagcttggt	gacatcagtg	ggccactggc	tactgggtgt	300
TTTTtccatc	ctgccctggt	ccctgctggc	accaggggaa	aaaggcccat	acaggagtct	360
gttccaggte	accagatcct	ggggtggccc	gtaaggattt	gaaggggaca	ggaggcgccc	420
ctttgccgag	gcccttcact	gtgtcaggca	cattgctggc	tgccgtgggt	gcacaatctt	480
agggaaaccct	cctgcctcct	ccgccttgct	gcttccttgc	aaagaaaatt	tcccactgca	540
gagggcagct	taattgctca	gcagtggctc	ttcagaatct	cacagatggg	ccaggcgtgg	600
tggctcatgc	ctgtaatccc	agcacttttg	gaggccgagg	cgggcagatc	atgaggtcag	660
gacatcgaga	ccatcctggc	taacatggtg	aaaccccatg	tctactaaaa	atacaaaaaa	720
attagccagg	cgtggtggtg	ggtgcctgta	gttccagcta	ctctggaggc	tgaggcagga	780
ggatcacttg	agcccaggag	gctgagggtg	cagtgaagtg	tgactgcacc	actgcactcc	840
agcctgggca	acagagcaag	accctatctc	aaaaaaaaaa	aaaaaaaaaa	aaaaactcga	900
g						901

## &lt;210&gt; 17

&lt;211&gt; 919

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

## &lt;400&gt; 17

ggcacagaga	aaagctagcc	actggtattt	tgTTTTgttt	aaaaaaaaaa	agaaagaaag	60
aaagaaagaa	aaacggaaag	gaacctagct	gcctgtatct	ttcattttta	aaatagcact	120
tgagttattt	tctgagtaat	ccaataaaga	actttttgatg	acagccagaa	tgtgttagaa	180
ctctggctga	acatttcac	tcctgtgagt	cagaagggtc	ttattttctcc	ctttgatggg	240
gccccttctt	ctttctggtg	ctctggaart	tgtttagrgg	aaagaattct	aattttaatt	300
aattgcgcrt	gagttaatct	cactcgcttt	tctgcttcca	ggcatcttag	gaaaaacmaa	360
tggTTTTtagt	aggcattccct	tatctactaa	tgctTTTTta	aaacaaacag	ggacattttt	420
attatagatt	tgattttttt	aatgaatgtt	tttaaaaata	tataaatagg	acaccaaaagc	480
ggcagggttt	TTTTtggggg	gaggggggtt	gtttyccaac	tcaagatggc	acattagtgg	540

ccagcaatat	tttttaactc	attccaacca	ggaagctttt	ttatacattg	cctaaatcta	600
cgccaaccag	aaaatagtct	catctctttt	tttctcaa	gagatccgtg	ttttatttta	660
gcattaaatt	agttacactg	tgatgactgg	cctattacct	gactcagctc	cctctacctt	720
gaaattgaca	tttttaaaaa	atgcaactaa	gtgggttaata	gtgtgtgacg	ctcaaagtta	780
atgtaaactg	gaaagggtgt	gtgtcgttgc	tttttgtgtt	ttgggttaggc	ttgggttttgt	840
tttttaattt	ttatactttc	taataaattt	gcagtttcat	tctttcaaaa	aaaaaaaaaa	900
aaaaaaaaaa	aaactcgag					919

&lt;210&gt; 18

&lt;211&gt; 1098

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 18

gcgagcccca	ggaggccct	gattccactg	ctgcaggagg	ctcagcctcg	aagcggatgg	60
cgctggtgct	ggaacgggtg	tgacgactc	tcctgggcct	ggaggaacac	ctgaatgccc	120
tggaccgggc	tgctggtgac	ggcgactgtg	gcaccacca	cagccgtgcg	gccagagcaa	180
tccaggagtg	gctgaaggag	ggcccacccc	ctgccagccc	tgcccagctg	ctctccaagt	240
tgtctgttct	gctcctggag	aagatgggag	gctcatctgg	ggcgctctat	ggcctgttcc	300
tgactgcggc	tgacagccc	ctgaaggcca	agaccagcct	cccagcctgg	tctgctgcca	360
tggatgccgg	cctggaagcc	atgcagaagt	atggcaaggc	tgctccaggg	gacaggacta	420
tgctggattc	tctgtgggca	gcggggcagg	agctccaagc	ctggaagagc	ccaggagctg	480
atctgttaca	agtcctgacc	aaagcagtca	agagtgccga	agctgcagcc	gaggccacca	540
agaatatgga	agctggagcc	ggaagagcca	gttatatcag	ctcagcacgg	ctggagcagc	600
cagaccccg	ggcgggtggca	gctgctgcca	tcctccgggc	catcttggag	gtcttgcaga	660
gctaggggtg	gtgactgcct	cccttggcct	cagctcctct	cactgctgtg	ctgaggtggc	720
ctttgtcact	tccttctgcc	ttccaaccct	caccttcccc	cggcctggcc	ccattggccc	780
accttctaag	ttgagcagga	aatcctccac	caagcttcca	gaactacaga	cagcaccag	840
agtgaagctg	agtgggtccc	catgcctctc	cagcatgccc	tttccctttg	caggaggggtg	900
gagtccttgg	gtcatgccct	cccctgccag	ctctgggctt	cagagataag	gcattttcct	960
tgtgcagcct	ttacctggca	atcctaattt	ggttttaaga	ctccctgtga	aatgctttcc	1020
gaaccttaac	cccagtgagc	gtgaaaaaga	aagttaataa	actataatac	atggaagcaa	1080
aaaaaaaaaa	aaaaaaaa					1098

&lt;210&gt; 19

&lt;211&gt; 1054

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 19

gaattcggca	cgagccccac	cactgtccac	tgaccagaaa	cctggctgca	gggccgagga	60
ctggtttggg	gactggaggg	ctggcagcag	cctgtcaccg	tgcgaccgtg	accacctggc	120
atgggcttcg	tgccctgctc	tcaggaagtg	ggtcaagccc	tgggaacctt	catccatgag	180
agctcgatcc	cgtatgaagg	gtgctgccgc	ccgtgccatc	tggcccgggg	gtgacttttt	240
gaactgttta	ttatatggtg	gatgatgatt	tcattctcac	tgctggacgc	tgttctgttc	300
agtgtgctct	ttggactaca	ttagtcccct	gtggagcagc	agggctggag	atctctgcag	360
tcccttcccc	gcccgccttg	ccagaaggcc	aaggaggcac	gtggagggcc	tccttcctgc	420
aattcttccc	tctccagagt	cagggagggc	tgcccagccc	tggcctcaca	gccgtcccag	480
atgttaggtg	agccactgag	ctctgtgttg	accttgaggg	gcctggctgg	gggccccag	540
gtcccatgcc	ttcttgggag	ggtggccgcc	aacgcctttc	ctgtgttatg	gcaacaggga	600
gtgggcatct	catctgcctg	tggtcagctc	tcagacggca	gggagcggag	ctgacgttgg	660
ctgtgcttgg	tcaccgctgc	catgccgcag	aggatgcgcc	tagctgggct	ggggccacac	720
gactattatg	ttggccttga	acggggactg	cagagccctc	agtttgtctc	ccttgttcct	780

ctgtggctga	ggtgggaggg	ggaggggtggg	gtaggtcccc	cagcaagaaa	gagggacagg	840
agcacccccag	gcaggaccaa	ggagtcggga	ggccccctgcc	ttctgtcctc	catggtgagg	900
gcacagatgt	ctccccagag	cccagcgctg	gcagaatgga	ttctgtcctc	ggctttgctt	960
ctgcggcttc	ggtggagaca	gttatggaat	aaaatgttcc	ttgcacccaa	aaaaaaaaaa	1020
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaa			1054

&lt;210&gt; 20

&lt;211&gt; 1667

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1659)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1664)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 20

tgggtggtgc	tggacctcgg	gcggggccgg	gcctcgtggc	gggaggaggc	agggcagggc	60
ctctgggacg	gggctggacg	gcttggtgac	ggaaacgagc	ccttgacgct	gtggcccgga	120
agtggagcgg	ctgtcgcagt	gcggctccgg	cagtggcagc	ggaggcctgt	gtttgcggcc	180
ttcggcaagc	gactgagatg	gcgagcgcaa	ctgcacctgc	agccgcagtc	cccaccctgg	240
cttcgccttt	ggagcagctc	cggcacttgg	cggaggagct	gcggttgctc	ctgcctcgag	300
tgcgggtcgg	cgaagccagg	agaccaccga	ggagttaa	cgagagatgt	tctggagaag	360
actcaatgag	gcagctgtga	ctgtgtcaag	ggaagccacg	actctgacca	tagtcttctc	420
tcagcttcca	ctgccgtctc	cacaggaaac	ccagaagtcc	tgtgaacaag	tccatgctgc	480
catcaaggca	tttattgcag	tgtactat	gcttccaaag	gatcagggga	tcaccctgag	540
aaagctggta	cggggcgcca	ccctggacat	cgtggatggc	atggctcagc	tcattggaagt	600
actttccgtc	actccaactc	agagccctga	gaacaatgac	cttatttcct	acaacagtgt	660
ctgggttgcg	tgccagcaga	tgcctcagat	accaagagat	aacaaagctg	cagctctttt	720
gatgctgacc	aagaatgtgg	atgtgtgtaa	ggatgcacat	gaagaaatgg	agcaggctgt	780
ggaagaatgt	gacccttact	ctggcctctt	gaatgatact	gaggagaaca	actctgacaa	840
ccacaatcat	gaggatgatg	tgttgggggt	tcccagcaat	caggacttgt	attggtcaga	900
ggacgatcaa	gagctcataa	tcccatgcct	tgcgctggtg	agagcatcca	aagcctgcct	960
gaagaaaatt	cggatgttag	tggcagagaa	tgggaagaag	gatcagggtg	cacagctgga	1020
tgacattgtg	gatatttctg	atgaaatcag	ccctagtgtg	gatgatttgg	ctctgagcat	1080
atatccacct	atgtgtcacc	tgaccgtgcg	aatcaattct	gcgaaacttg	tatctgtttt	1140
aaagaaggca	cttgaaatta	caaaagcaag	tcattgtgacc	cctcagccag	aagatagttg	1200
gatcccttta	cttattaatg	ccattgatca	ttgcatgaat	agaatcaagg	agctcactca	1260
gagtgaactt	gaattatgac	ttttcaggct	catttgtact	ctcttcccct	ctcatcgtca	1320
tggtcaggct	ctgataacctg	cttttaaaat	ggagctagaa	tgcttgctgg	attgaaaggg	1380
agtgcctatc	tatatattagc	aagagacact	attaccaag	attgttgggt	aggccagatt	1440
gacacctatt	tataaaccat	atgcgtatat	ttttctgtgc	tatatatgaa	aaataattgc	1500
atgatttctc	attcctgagt	catttctcag	agattcctag	gaaagctgcc	ttattctctt	1560
tttgagtaaa	agtatgttgt	tttcattgta	aagatgttga	tgggtctcaat	aaaatgctaa	1620
cttgccagtg	attaaaaaaa	aaaaaaaagg	gcggccgcnc	tagnggg		1667

&lt;210&gt; 21

&lt;211&gt; 1759

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (367)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (896)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1747)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1748)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1754)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 21

gaaaacattt	aacaattcaa	attatatgca	ccttttacct	agttgaaaaa	aatacacatt	60
cctgtkttca	cattatagca	actgattaag	ctgaagctgt	aagtcatttt	ttatagatga	120
gtgatccgca	tctccatcaa	ttagaacact	ggaaaagatg	ttttataaaa	gaggtattta	180
atthttgtttg	taggattaac	tcattgcaaat	aataaaaaag	atatcctgtt	ggttcaatag	240
tacactgtct	cctttaagga	aggaagcgtg	atgaatgaat	gatgtgtaga	cttgagggat	300
gactattaaa	ggggacgtag	gatgaagaga	aagaacctac	agatgacaat	gaatgtaaac	360
ttatttntct	tcattgtgtaa	gcagtgtgct	cgctgggtgat	atccagatcc	taacaagatt	420
acttggttag	ctggtttagga	ccagtaactg	gattgctgacc	actatgataa	tatttttgaa	480
caaatgttaa	tgcttgatgc	agaattgtaa	agcagcatct	ggttcctata	tagccttaag	540
gattaatttt	agtgatcctc	aaggaattaa	atagggaatt	tcagaaatgt	agactgcaaa	600
ggcagtatac	aggaaaagg	ggagtgggtt	ttgtttatga	gggtgtctga	aaactaaaat	660
tgagcgggat	atcatgggtat	agttggacag	tattggctct	tcacactttg	gccatattgt	720
ataatggagc	ttttaccaaa	gatgtatgag	aagtgtgaaga	ctataaaaaa	atgaactatt	780
caaagtaaaa	ctcttaacaa	acattttact	taaagcagat	gcaaaaagggt	attctcatgt	840
aggctcctgt	tggtgcagag	ggattttttt	gatttcagga	tacaactaaa	gtacgnagtt	900
ctcagtttca	cttttagtaga	aagagctcta	gaaatgaggc	tgataaacac	atctaagaac	960
actgggttgc	ttctaaaatt	tccaaagctc	caccataaat	gtaattttta	gtgtttcaaa	1020
tgattgcatt	ttaaagtata	taaatatggg	ttatccaata	tcaatgctat	agtaacatcc	1080
tgaacaaaaa	caagcacaaa	ggtataaatg	cctaaactgg	aggaaacttg	aaaccctcat	1140
gttaaattct	aaatgtagta	tttctaactt	gtgaagacag	attggtaggc	agccattttt	1200
ttgtgtctta	aaataactgg	gggcatagtt	aaaattttat	acatcaagtg	attgctatta	1260
ttgaatggtg	caggtgagat	gtggttattt	ttagttttatt	tgaaatggtt	gactggaaag	1320
gggggagggg	gaagcaaata	tttgaaattt	ggaaaaccct	aaaccttttg	gtaagaaatt	1380
gtaattttca	cttaaaattt	tctttaagga	tataagaggt	ttataattga	tgtagttaaa	1440
ttgaacaata	accattgggtg	actggagcag	gtaattatag	cctgcagaaa	aaattatcta	1500
agaattttta	aaataagatc	ctgaagttgt	ttaattgcat	ccatttctgt	atttatgtga	1560

at ttataaac	tgcagtaagt	tttgaatgag	gttaatccttg	tttaatatataa	gtaaatgagt	1620
ctgtagactg	tgatctcccc	aaactaaaaa	gtacagtact	tggaattgtg	ttctttatgg	1680
ttgtagtggt	ggtaaagcac	taatatgcag	aaaataaagg	aattacacag	tgmaaaaaaa	1740
aaaaatnnaa	aaanaaaaaa					1759

&lt;210&gt; 22

&lt;211&gt; 658

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 22

tcccttgctt	gatgcaacac	tgttccaaaa	gagggtggag	agtcctgcat	catagccacc	60
aaatagttag	gaccggggct	gaggccacac	agataggggc	ctgatggagg	agaggataga	120
agttgaatgt	cctgatggcc	atgagcagtt	gagtggcaca	gcctggcacc	aggagcaggt	180
ccttgtaatg	gagttagtgt	ccagtcagct	gagctccacc	ctgatgccag	tggtgagtgt	240
tcatcggcct	gttaccgtta	gtacctgtgt	tccctcacca	ggccatcctg	tcaaacgagc	300
ccattttctc	caaagtggaa	tctgaccaag	catgagagag	atctgtctat	gggaccagtg	360
gcttggtatc	tgccacaccc	ataaatcctt	gtgtgttaac	ttctagctgc	ctggggctgg	420
ccctgctcag	acaaatctgc	tccctgggca	tctttggcca	ggcttctgcc	ctctgcagct	480
gggaccctc	acttgctgc	catgctctgc	tcggcttcag	tctccaggag	acagtgggtca	540
cctctccctg	ccaatacttt	ttttaatttg	catttttttt	catttggggc	caaaagtcca	600
gtgaaattgt	aagcttcaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	658

&lt;210&gt; 23

&lt;211&gt; 1463

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 23

ggggggagccc	gcggctgctg	ggagctgcgg	cgctggccct	ggggggagcc	ctggggctgt	60
accacacggc	gcgggtggac	ctgcgcgcc	aggacctcca	cgcagagcgc	tcagccgcgc	120
agctctccct	gtccagccgc	ctgcagctga	ccctgtacca	gtacaagacg	tgtcccttct	180
gcagcaaggt	ccgagccttc	ctcgacttcc	atgccctgcc	ctaccagggtg	gtggagggtga	240
accctgtgcg	cagggctgag	atcaagttct	cctcctacag	aaagggtgcc	atcctggtgg	300
cccaggaagg	agaaagctcg	caacaactaa	atgactcctc	tgtcatcatc	agcgcctca	360
agacctacct	ggtgtcgggg	cagcccctgg	aagagatcat	cacctactac	ccagccatga	420
aggctgtgaa	cgagcagggc	aaggagggtga	ccgagttcgg	caataagtac	tggctcatgc	480
tcaacgagaa	ggaggcccag	caagtgtatg	gtgggaagga	ggccaggacg	gaggagatga	540
agtggcggca	gtgggcggac	gactggctgg	tgcacctgat	ctcccccaat	gtgtaccgca	600
cgcccaccga	ggctctggcg	tcctttgact	acattgtccg	cgagggcaag	ttcgagccg	660
tggaggggtgc	cgtggccaag	tacatgggtg	cagcggccat	gtacctcatc	agcaagcgac	720
tcaagagcag	gcaccgcctc	caggacaacg	tgcgcgagga	cctctatgag	gctgctgaca	780
agtgggtggc	tgctgtgggc	aaggaccggc	ccttcatggg	gggccagaag	ccgaatctcg	840
ctgatttggc	ggtgtatggc	gtgctgcgtg	tgatggaggg	gctggatgca	ttcgatgacc	900
tgatgcagca	cacgcacatc	cagccctggg	acctgcggkt	ggagagggcc	atcaccgagg	960
cctccccagc	gcaactgaatg	tccccgcgca	gagcagaggg	aaggcagcgg	aagacgccag	1020
ctgccagggc	ctggggccac	tggggccagcg	cctggcgata	ctgggtgggg	gcaggatcat	1080
tctgcccctt	gtccacgcac	ccccaccagc	cctctcgctt	ctaacacagg	gcacctgctg	1140
gggctcaggg	atgttagagg	cgagttccag	ccctgccact	gccctggggc	gacccctccc	1200
tgtccctgcc	tccctgctct	gcgcgcctc	ttcctggacc	ctcagtggct	gtcccatggc	1260
tacatcctgt	gggtgggggc	cctcgacagg	acagcaggac	ggttttgttt	cagtggaatc	1320
ccatccctgg	gttcccctgg	ttcccactct	tcccaagcct	cccgggactg	ggacatgttt	1380
gcaataaagg	aaaggtttgt	ggcgccctgtc	atggcaggca	tctcatggaa	aaaaaaaaaa	1440

aaaaaaaaaa aaaaaaaaaa aaa

1463

&lt;210&gt; 24

&lt;211&gt; 475

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 24

ggcacgagct	gacgcagcct	atcgtggtca	ccgtgccgcg	gccgcccccc	aggccgccca	60
agagtgtccc	cggccgtgca	gtgcgccctg	agcctcccgc	gccggcccccc	gcggccctgg	120
aaccgcgcgc	ggtggtggcg	ctggtggttg	cagccttcgt	gctgggcgcc	gcgctggccg	180
ccgggctggg	tctcgtctgt	gcgcactcag	cgccccacgc	ccctggcccc	cccgcgagag	240
cctcgcccag	cggccccag	cccaggaggt	cccagtgagg	aagggatggt	gcgcccccaa	300
catggtccgg	agatacaccc	agctaccaat	tcgggaccag	gaccaacagg	accggacccg	360
cctccctgga	cctcggacct	gatgaggcca	cgacccctgc	gcttctctcc	tccccctgtc	420
cctcccacct	gtgctcaaaa	taaacctctg	gactgacaaa	aaaaaaaaaa	aaaaa	475

&lt;210&gt; 25

&lt;211&gt; 520

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 25

ggcacgagcg	gagtttgagc	cccggaggca	gagcggctgc	catggccaag	tacctggccc	60
agatcattgt	gatgggcgtg	caggtggtgg	gcagggcctt	tgcacgggcc	ttgcggcagg	120
agtttgacgc	cagccggggc	gcagctgatg	cccgaggacg	cgctggacac	cggctctgcag	180
ccgcttccaa	cctctccggc	ctcagcctcc	aggaggcaca	gcagattctc	aacgtgtcca	240
agctgagccc	tgaggaggtc	cagaagaact	atgaacactt	atttaagggt	aatgataaat	300
ccgtgggtgg	ctccttctac	ctgcagtcaa	aggtggtccg	cgcaaaggag	cgcctggatg	360
aggaactcaa	aatccaggcc	caggaggaca	gagaaaaagg	gcagatgccc	catacgtgac	420
tgctcggctc	cccccgccca	ccccgcgcgc	tctaatttat	agcttggtaa	taaatttctt	480
ttctgcaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa			520

&lt;210&gt; 26

&lt;211&gt; 1054

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (89)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (149)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 26

cgcaattaat	gtgagttagc	ttcactcatt	aggcacccca	ggcctttaca	ctttatgctt	60
ccggctcgta	tgttggtggt	aattgtganc	ggataacaat	ttcacacagg	aaacagctat	120
gaccatgatt	acgccaagct	ctaatacgnc	tcactatagg	gaaagctggt	acgcctgcag	180

gtaccggtcc	ggaattcccg	ggtcgaacca	cgcgtccgag	tgacactact	gacttggtc	240
taccacctga	aatgccgatt	ttgattgatt	tccatgctct	gaaagacatc	cttgggcccc	300
cgatgtatga	aatggaggtg	attcatttctt	tttatttctt	tttgctccag	tcaatgaaag	360
gaacacttta	ttgaggcccc	agggccgtag	ggcctgggca	ggaggctgcc	ctttggggaa	420
ggaatagcct	tattcgacct	tcttttttggg	acgcagggtg	ttggtgtggc	cgcacttctt	480
gcagcagttg	actgcatggg	ggcgcaggcg	agcacagctc	ttgtggcaca	tcacttctt	540
gcagttgtat	ttctgggcaa	ggtggcagag	ggaaggctcc	gtaatgccac	ctcacaggca	600
cagcatcagg	cgcagggtgg	actctttctg	gatgtttag	tctaagagt	tgtggccatc	660
cttcagctgt	ttgccctcaa	atatcagaca	ctgctggtca	ggtaagatgc	cctacctgtc	720
ttgaattttg	gctttgacat	tctcagtggc	atcactgggc	tcgacctcaa	gggtgatggt	780
ctggcctgtg	agggctcttca	caaagatcca	catctcagcg	tctgcagctt	ggccagtctc	840
actccattct	catttttttg	ttggtactca	ctggtgtact	cagggtggtg	cttaacagag	900
aagtaaaatt	ggatgtttcc	agaggctgaa	ttttgcctta	agatggaamc	tttatttcta	960
aaaaaaaaaa	aaaaaaaaaa	aaaagggcgg	ccgctctaga	ggatccaagc	ttacgtacgc	1020
gtgcatgcga	agtcataact	cgtctatag	aatg			1054

&lt;210&gt; 27

&lt;211&gt; 744

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (271)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 27

ggccccagga	cagccggagc	tcggcctgcg	gtggggggcg	ttcctgtgga	cgtgcgcct	60
cccatcgccc	gggtgcctgt	ccttgcccag	cgccaccagg	ctggaggcgg	agtgggagga	120
gctggagcca	ggcccgtaa	ttcgcaggca	ggggtgggtg	tgggacgggg	ctgcttctct	180
acacakctc	yacgctggcc	ttcaccttca	cccctgcac	gtcggtgacc	ctgggaccct	240
ccaggcagcg	tggcctgtgg	caccgtgagg	nttgggaccc	accgaggcgc	agaggcggcc	300
cgaatgcagc	cctggttcag	gcccggagag	ggtttgcggg	tagttgcacg	gacaattcgg	360
cggggtgctg	cctggttctg	ccattagccc	aggaggaggt	cgtgggacgg	ggagggtggg	420
atggacggcg	gacaggcagt	ccccacgctg	ctgggtggcg	ccgggcttgg	tggggtcttc	480
cactgtgtgc	ccttctcgcc	gaggccggtc	ccccgggtgt	ggggtgccct	gctgcggact	540
cctccgcgag	ccccatcgtc	gcgcctgtgg	acgcctaggc	aagagcggcc	ctctgcagcc	600
aagagaaata	aaatactggc	ttccagataa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	660
aaaaaaaaaa	aaaaaaaaaa	ggsgggccgc	tctagaggat	ccaagcttac	gtacgcgtgc	720
atgcgaagtc	ataactcggc	tata				744

&lt;210&gt; 28

&lt;211&gt; 1463

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 28

ccacgcgtcc	gtgtatttct	aactgcctga	gtcacacaga	atagggtaag	agcctgaccc	60
cattctgtaa	atcagaaagc	aaggatggag	accctttctt	gctgctatta	ttggctctct	120
ttgaggaagt	tggagggtta	ggaaggaaact	tgtttgtttc	cgtatacgac	tccttcttct	180
ctctagtcca	gtcttcagcc	agtccagcgc	tctcttccac	acttcagagc	cccttcagag	240
aaagcattag	caggaatgag	acaaggcaga	gctgcagtcg	cccctgaggc	ttccacacat	300
ctttctgaat	attatttttc	aagtaacaag	ggcagggaca	gcggaacag	ctgcccaccc	360

ccccatccc	agcagctcag	ctaagccctg	atgagaatga	agccacagga	gttgtctgag	420
gtgaacccag	ccgctcagcc	acacatggaa	gccattgcct	ttgcacatag	ttcttgggtt	480
ctttttccta	aaaaggtaag	gagctgaggt	gtgtggtttt	ttaatatataa	gaatatataa	540
tggaaaacac	acgactgacg	ctcaggcatc	ttcccctact	ccccaacaga	tccccagaag	600
acagcgtgga	aggcagtgtg	gacagtaaat	cgggcttcag	ttctatagcc	aagaagagat	660
cagctgctga	aaccaccagt	gggtacccca	ggccacctgc	ctttgaactt	ggggatttgc	720
catgtctgat	cttgtcacat	acttgctttt	ttacaagatg	aactctttgt	atttatgatt	780
tggggggcaa	tgaaagggtg	aatgcaggaa	ctgctgctgc	cgagctcgct	ggtcacatgg	840
gggtgccagg	cgggatttctg	gaaaaccagt	gcacttaaac	tgatcctgaa	gagagctgtc	900
ccagcactct	ggccaccagg	agggccagat	tccccagaaa	ctaccttttg	cccaaagaac	960
atgctcagta	tttggggcat	ttcctcccac	aaaccctgac	tgcttctgtt	acctcagggc	1020
cttggtacct	ggatactgcc	acagaattgg	ggcgggtggg	ggaggggcct	atttttaaat	1080
aaaataactg	ttcaaagttg	ggggtttttt	taaaaaatta	agaaaaagga	aagctattct	1140
gtattgcacc	ttttcacaat	ttaatacatt	ttcttacatt	ttcctgtgat	tttcgaaact	1200
aaaccattgt	gtgcctgtag	tgtcctggtt	gagctgccgc	tcagcagctt	cctcgggggg	1260
atttggaaca	cctgtgtctg	tcgccgcact	gcctgtggga	ggggcccaga	gggctgctgg	1320
gactggcgct	tgtacacact	tgtttggcct	tttctgtagt	tgatgctgta	aactctatgg	1380
cttttttaaaa	acgatttcat	gtttttatatt	agtattggaa	atccaataca	cttttttaaat	1440
ccaatcaaaa	aaaaaaaaaaa	agg				1463

&lt;210&gt; 29

&lt;211&gt; 214

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 29

ggcacgagct	cgtgccgaat	tcggcacgag	aggacggagg	cttttggacc	ctcggacccc	60
atcccactca	gccaagtgtc	tttctgtgtc	tggggggagg	aggggatgat	atccgtgtgg	120
ttcgatgtat	tatttttaag	ctccgtgagt	gcgtgggtca	gtgtctgcat	gaagtggaat	180
aaactgcccc	ccgccaaaaa	aaaaaaaaaa	aaaa			214

&lt;210&gt; 30

&lt;211&gt; 1159

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 30

gtgctatcct	actatgctgt	tctttggtaa	tggaataaat	tgacccaagg	accgaatttc	60
atttggaattt	caaattgtcc	agagtggaaa	agccttcaag	atgacatgat	gaattactca	120
gttcatctga	tttctggtec	ctcctttctc	gacaactata	atactaacc	ttttctcagg	180
ataactgtct	acacctggca	gttttctctg	acgtgctgtt	cactcacatc	cctaccttgc	240
atggtaatat	aaaggactag	gaagcagtca	tacttccagg	aatgcttgg	attcatgtgg	300
acattcagga	agcttattct	catataatac	taatctaaac	agtactagaa	attacagtgc	360
caagagccac	caggaggccc	agccaataag	catagatact	atatgggtatc	atgggaccca	420
tctattttttt	accagtggac	tacaggatta	cttgagagtt	atcagggctg	cctaacagac	480
caggagatct	gggggttgca	ccagggaatc	gccatatttg	accagcatgt	tttaaaagct	540
cttggttagga	ttagttggtt	ctaaggatcc	ctctagggac	ctcattattt	caagaggaa	600
ccaaagtcca	gcctcctaca	tagatgctgc	cccacgaagg	accacaaaa	ctaacctagt	660
tcagggttct	caggcaggca	gttctgcttc	agcttagagc	agaaccata	aaataactca	720
gtactgggat	aggcaagcat	gtgtgtttac	tgtggatttg	tccctgaagg	ctcctttggg	780
tgagaacatg	tgaaccaggc	accctgggtt	gtttggagca	ttgctgcccc	gaagcttcta	840
tgggataggt	ggtgcttggg	attgatgtgt	tgtggccatg	cagccctccc	tgaggattga	900
cttctgcact	aatccagtga	aggaggctgt	gtcaaaagaa	gggctcagaa	gccctctttt	960



cagaggcaat	gattcctgtc	agtatgaggt	cccttagtta	ctaaaaaggg	acatgattta	1020
actccagttt	gatgaacctc	ctccgagttt	actttattgt	cttcaaactc	tttgttttct	1080
tcctttttgt	gagatttgtg	ggttttgtgc	cttataaatg	gaaatgtatg	aacacaaaaa	1140
aaaaaaaaaa	aaactcgag					1159

&lt;210&gt; 31

&lt;211&gt; 3116

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (1342)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 31

atccacccgc	cttcagcctc	ccaacgtgct	gggattacag	ccgtgaagcc	gccgcgcctg	60
gcctagactt	gtgtgttcta	aacagggttaa	gtagcagggt	gggtttttat	gatagtgcaa	120
ggaatgactc	atgcctctga	gcttctaaac	tgaagctgct	gtaactaaag	gaatctgaaa	180
agaacaaccc	tgaagcagag	gcctttattg	tcttggttgc	cagtagtacc	ttgttttgcc	240
atgtagcaga	caacacacaa	aataatgcag	ttgtgggtgtg	ccatgctatg	tgcacagccc	300
cttggattac	tttgttttta	aaagcatcag	agttgggggt	actttaggga	aacctttgct	360
taccttgttt	tgccagtgat	aagagcagtg	ggttggaggg	cacttgGCCA	gttttctggt	420
cagcttttca	gtgaatgtac	ccctttaagg	ttcagactta	aacttcctta	aaaagtggcg	480
ttgttcatag	aatcggttga	ctcattaatg	aatcgttcaa	ctccactcac	tgaagcccag	540
acctccgtgc	ccaggGCCCA	atctcgtcag	gctgccagag	aaagtgggtg	ctgctcatac	600
tggtctcaca	gtctaagtaa	gtgtctgtga	tgctcccaag	caaaggaaat	gcaagctctg	660
gaaattcggt	aatgtatttg	atgtcttagt	gttttagtga	ctagggagac	cattaactag	720
tttatcatta	accacttatc	agtgtattga	tgttaaagca	tttccctggt	agctaaaaga	780
ggcctgttca	tacaagccaa	ctggtatata	cgtgtgggtc	atccatcatc	tgctgcacat	840
agcagactag	aattctggga	accctgtgca	attcagctctg	ctctcccttg	tggaccctgg	900
taaagaaaag	cctcagctca	tagtgaacac	agcagacctc	gaaatgtagc	agcagcctac	960
tgagtagctt	tcatttactg	atcatctgct	gtgactgtgg	ccctgtctgg	aggttcctag	1020
gttttgagat	ttagagcaat	gcattctgga	gacagaacca	gcagaacagc	catttttcaa	1080
tttttcttta	aatcagttat	ccatcaggca	gataactgct	gtattcatga	atcttgagag	1140
tgttcctgag	acagaattaa	tgggtcatttg	ggaaaactat	cgccatgggt	tcccatctgt	1200
ggttttcctc	taaaagcctt	ggagatttagc	ccttccttgc	cagtgagaac	ggtgaccgcc	1260
tcctgctctg	cacggtctgc	ggcagttgcc	cttctgggtta	ggtgtgtcag	gttggtttat	1320
tttggggttca	ggcctggcgt	anaccacaa	gtggcagaca	tatcacaaga	gtccccagac	1380
tctgcctaga	aacagtgttt	gccctttggc	cagtgcagtg	gttcatcccc	gcccagtttg	1440
agccatgagt	ggagtttcca	acagagggag	gaatgtgtgc	cttggtcaag	gagggcacga	1500
cccttagggc	tttttcaacc	agatttagct	gaagggttg	acacctttga	attacagcag	1560
ttgactcaga	gtgcaagaag	tctggccatt	ttggaaagca	aggtttcctt	tcagccctgt	1620
ctactgacca	ataccccgac	tcaccttggtg	tggcgcactt	cagaatcaga	tataacctaga	1680
gtataacctgt	ggtttggttt	tataattaat	cagctcggtta	cttcagccca	tgaaaatggc	1740
atccagggtc	gccaggagat	tcagagctca	aaacaaggcg	agcttgagtt	ctgcactcca	1800
gatgtgtgcc	aaaactagta	aaacttaacg	gacttacaac	cttggtcagtt	tttttaataa	1860
ggcagggtata	ctctgttttt	cacactaaac	atatgaatgc	agcactgctg	cctcagctca	1920
gcttcgtgcc	tgggttcccc	actgggtctgg	gaagactggt	gtgctccata	gagcagtgca	1980
catctgaccc	agagggtggg	tgttcataac	tgctacttgc	tctgctctac	catgtttaaa	2040
gaaatatttg	gatgttaaat	taactcacta	tgggttttca	cctgggaagg	aaacaaatta	2100
cgtactagag	ggcattgatt	ggttaaaaac	ttgtgtatcc	cgggaaggac	ctgcggtaca	2160
ggagtcagcc	atgtctgtgc	tgtgtggaac	cacctgatga	catggttaac	gaggaagacg	2220
atgtgttgac	cggctgccgt	ttgaggactt	tggtcaccca	gactagacac	cttctgtgct	2280

catgtttgga	aagctgaaag	ggaaggacag	ctgtgccctc	ctgggagctc	atgtgtccct	2340
ggcgtgtgc	tagctttcct	ttacagctgt	ttacagacaa	ggcaggcctg	aggcagatgg	2400
ccactgctct	tgtgatgttt	gctcagagga	atatgaacat	tttatttttg	aaaagggatg	2460
atgtggtttt	ttgccagggtg	tttataatta	atcctttaat	attatggtta	ttaacctctt	2520
aaacatgaat	gaattcttga	ttgttttaac	acagtaccta	agactaatgc	tttctgtgga	2580
caccactgag	ctctgcctca	actccaccct	ctgcgaccgg	aggactatgc	ccctagtaac	2640
tgctgtcggg	gtggacgctg	tgctggttct	gttttctaaa	ggagcagaag	gacaggctctc	2700
tgagacagga	tcgttgctcc	tacaggagga	acagtggcct	tgcttcttag	acggctcttca	2760
ctgtgtgttt	taaaacaaca	acaacaacaa	caacaacata	aaactctttt	gacctgtaac	2820
ttaaagatca	taaacttcag	gcaataatat	tttctgtgta	agcttttaaa	attatttttg	2880
gggatcatag	cttggtttat	tttgtgctat	aaaattaaca	gtattaaatg	acttatattc	2940
ttagaataca	tcgagtgtct	tttcttaaca	gattagtgcc	tttttatttt	tgtattccgt	3000
tttacgttac	tggtcccagc	atcaaaaccc	ttgtttccat	ggcctgtttg	tatattgtct	3060
caataaaaact	tgcatcagcc	ggtggtggcg	gcaaaaaaaaa	aaaaaaaaaaa	aaaaaa	3116

&lt;210&gt; 32

&lt;211&gt; 922

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 32

tgcttttcag	gctcttagaa	gccatagatt	tggacaagcc	cagcaagatg	ggtgtccttc	60
caggcctctt	cccctttcct	ccatctctgg	caacagttct	tgggggtttg	caattgtttg	120
gatttttttt	ctttctgcag	ttgtgtgtat	gtgtgtttgt	gtgaagaaaa	acagactctg	180
tccaggtaga	aatgggtgagg	agggggaaga	gaattacatt	tccagggtca	gaaacttggc	240
aacagttttc	ctagagtgc	tcagacacac	cacagtaaca	actctcgtctg	caattttatt	300
ttaatttgag	aaataaagat	ttcctccaag	ccacatgagg	actctggcac	ccacccacaa	360
agcaagacct	gtattttataa	gccgagggct	cagggagcct	aactgcggga	cccgtcaggg	420
ccccgtgacc	catccccgtc	cccaccccc	cctccaccgc	tgggcccac	agtgtgtgtt	480
ggggggatgc	ttggcagctg	ggggtgagga	gacaacaaac	ctcgggaact	ggagccagag	540
ctgcggcctg	actgacgcct	tttgatgctc	acgggaaatt	tctgccagag	atctcagccc	600
caggctgggt	gtttctacaa	atctctctca	aatgtattat	tttggtgaca	aaaatgaagg	660
agctttgtaa	attttttttaa	aattatgaat	catatcaagt	agttgtttac	atttcttgaa	720
aaaataggaa	ctcgggcagc	agaatcagat	tggcagaatc	tttagactac	acaggcaata	780
atcaagtctg	ctgtttttggc	ctttcgtagt	agaagtgggt	gtagtgttta	gatattctgtt	840
tggtcttgct	tcttgatttg	catttttttc	aataaacaac	aacaaaaaga	aaaaaaaaaa	900
aaaaaaaaaa	aaaaaaaaaa	aa				922

&lt;210&gt; 33

&lt;211&gt; 2631

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (2619)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 33

ggcagagcca	cagtgataga	atacgtgaag	ccctcagacc	tcaaaaagga	catgaacgag	60
accttcaggg	agaagttccc	ccatgtcaaa	ctgacgctga	gcaaaatcag	gagcttaaag	120
cgkgagatgc	ggagcctgtc	ggaggagtgc	agcctggagc	ccgtgacggg	ggccatggcc	180
tacgtgtact	ttgagaagct	ggtcctgcag	ggcaagctca	gcaaacagaa	ccgcaagctg	240

tgcgctggcg	cctgcgtgct	gctggctgcc	aagatcagca	gtgacctgcg	caagagcggc	300
gtgaagcagc	tcacgcataa	gttagaagaa	aggtttcgat	tcaacaggcg	cgacctgata	360
gggttttagt	tcacagtkct	cgtggccttg	gagctggccc	tgtatcttcc	cgagaaccaa	420
gtgttacctc	attacaggcg	cctcacccag	cagttctagc	agaggcccca	cagaaggctc	480
agggcaccca	gggtgcactg	ccggcctggg	agggtgccca	ctgaagcccc	gcgcctcctc	540
ctgccagcac	ccccagcacc	tgctagcagg	aggcacctgg	cctycgctgg	tgacgctttc	600
ctttttgcct	ctttgccatt	tccttggaag	gagacgtcgc	tttcatcccc	aagtgcaccg	660
tccctccgag	gggattttctg	agaattctcc	tgcattttta	cataaactaa	atgkgagggt	720
tgttactggg	atttttttca	cgtgcctgag	accagcctgg	taccaggacc	ttttgttcac	780
agcgtgcagc	agcgagccgg	ctgcagtgyg	tctcccctgg	cctcgccttc	tgcaaaccac	840
cgcagccacc	acagcgtcag	gggtggagatc	tgggttttcta	gacctcactg	aacacactga	900
aatggctgag	tttaacttat	ttaggcattc	atcttgagag	tgtgggtttt	cgggttcccc	960
agcagcatct	cccgcaccca	actgtgccgc	tggctccctg	ccacctgaag	ccgagctcct	1020
ccagagcttt	ctccgcccac	ctcactgcat	cccaagtggg	gcttttggtg	tccagttagg	1080
ccagcgggag	cagtctcctg	atcttatttt	atctcattct	tggactcttg	gacctctctg	1140
ttcttcaagc	atcgtgtcac	tgtgaaatcc	taacgcccct	gtgtcctaca	gaccgacggc	1200
acaacagaca	gctgcccata	ccatgccatg	ctctaccctc	tgctctctac	caggagacac	1260
tctgggcctc	caggacaatt	gctgcttgcc	ggctcttatt	tttctaagca	atattgtgat	1320
ggagaaaaat	aacatattta	ttgggatttg	gttttttggg	tctttttttt	ttaagggaac	1380
aaaaaatggt	taaatgagggt	ctgctgaagt	tgacttgaaa	acacactkga	ccctcaggca	1440
ggagggcact	gaccacaccc	cacacaacct	caaagggtca	gtgcgtcagt	gccttttctt	1500
ctgaggcagg	aaacagggtgc	catcttgggc	acctcggcca	gggcagccca	ccatgctaaa	1560
aggaccccaa	atgggtggcg	ttgtcccttc	tgtgcaggcc	agcaggggcc	catctctagt	1620
ttttccacgt	ctgtctgaag	ttcttgcaac	aaattctgca	tgggtccagcg	ctccagctag	1680
ctgcctcatc	aaaaacactg	aataaccaag	gactgctgag	tttttcttca	tgggggggtca	1740
gctgggtctca	aaactggcca	ctgcctcagc	caccaagctt	tttctacca	ctaccttata	1800
aacctgcctg	gccctggagg	ggytctggga	cgactttgtc	tctagcccat	taatacaata	1860
catcctatgc	tttctgtgca	gactgggtgct	tccgcagaaa	ggagatgcca	attctgctat	1920
cacagaactc	caccagcaac	tccacccgac	cccagcagtg	gtgcaggaca	gctgccagca	1980
cccacctggc	cctcctcctt	ttccacagcc	actcactggg	gccacaaaaa	cccagagatg	2040
gcaggtgtgt	gggacagact	ggaggatgag	gacaaaccaa	agcctttgtt	ctttcttatt	2100
gtggagcgtc	cccttccatc	aagcagcctg	ccctcaagcc	aaggatcatc	ccctggaacc	2160
cagtatgcac	ccagagggga	ctcarcttca	aagctgctcc	accatgctgg	gtcccaggca	2220
gctttcctct	gaaaagcaac	ctctcctgcc	accagatcc	catctcaaga	gctcgcccat	2280
gttacgagca	tgtaaaggac	tgacttccta	gtaactgttg	cagtttacia	cctgccctcc	2340
agggacacgt	ctccatggtt	ttttctacac	atgaatgcga	gaaatggctg	ataagcacag	2400
taattttaaga	ttgcgtattc	atgtagttag	attttagtcat	taaccaaccc	ggtctttgtg	2460
atgtgtgaag	ccttcccttc	agttatgtcc	catttttatg	atgccaaaag	ctgccactgt	2520
gtgggtattcg	aggattattg	caacaaagcc	agtagttaaa	ccaaaaaaaa	aaaaaaaaaa	2580
ctcagagggg	ggcccgttac	ccaattcgcc	agagagtgn	aatggcatca	a	2631

&lt;210&gt; 34

&lt;211&gt; 385

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 34

attcggcacg	agtaataatc	aacaccctcc	tagccttact	actaataatt	attacatttt	60
gactaccaca	actcaacggc	tacatagaaa	aatccacccc	ttacgagtgc	ggcttcgacc	120
ctatatcccc	cgcccgcgtc	cctttctcca	taaaattctt	cttagtagct	attaccttct	180
tattatttga	tctagaaatt	gccctccttt	tacccttacc	atgagcccta	caaacaacta	240
acctgccact	aatagttatg	tcacccctct	tattaatcat	catcctagcc	ctaagtctgg	300
cctatgagtg	actacaaaaa	ggattagact	gaaccgaata	aaaaaaaaaa	aaaaaaaaact	360
cgaggggggg	sccggtaccc	attgg				385

<210> 35  
 <211> 2146  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (20)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (38)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (2125)  
 <223> n equals a,t,g, or c

<400> 35  
 cgtgcccagc ctattcatgn actcaactgg gatctggnc cactgcactg ggtgttttgt 60  
 tctctgatgt gcacattgat gaatgcatat tgggtacaagg gaaatgaatc tgttttcagc 120  
 catttaaaaa attgtgttgt aaaggcaaac aaaaatagct tgaacattta aaaattactg 180  
 tttaatatgt tttgactgat tttctaagaa ttatgtgaac aatgatgagt ccagggcagc 240  
 agccatcaca aatatgacag tccagcaagg cccatggagg tccctttgga agggacttgc 300  
 acagctgatt ttctttttcc agtgttggtt ttacgggtga aacgtggctt ttgaaatact 360  
 ttagttgttt ggcttagagc tcccactttc cccctaacat agtgcctgtc tgttttttaa 420  
 aacattttac ccttcggcac tattttctgc tattcccatt aatattatgt aacataaacc 480  
 acttgctact gatttgggca tattacagac tacaattact agttgctttg aatgcattgg 540  
 cactcttaat tattgcaaag gagaattgaa aagtagtctt gtggtcctag ttaatttagc 600  
 tttgggcaaa ctaagtggct tttcttctgt cctcttctca aaaggccttt taagcagcag 660  
 agactgggct aaatactgga actgaattcc acttttctgc ctgctcattt gctgccaag 720  
 accacttggg gctccttgtc tagaatgctc taagacgaaa tttggccaga cacagggttc 780  
 cttgattgyt ctgttcaact aatagagaac ytacattttg ctttcttaat atcgactttc 840  
 ttttgtacca ggtttttctt tctttctttt ttgccttagg agcttagctc aaagttcagc 900  
 taacctttga cactactagt gtcatttagt aacagtaatc tttgacacta ctagtgtcat 960  
 ttaggagctt agctcaaagt tcagctaacc tttgacactg ctagtgcgtt caagaaagg 1020  
 tagggaatga agtaccctta atgcagggtc agcaaacttt tcctgtaaaa ggccagatag 1080  
 gaagtatttt tggctttttg ggctgtttgc ttttgctact actactgtgg gaatgaccat 1140  
 ggtcatggct gtgttccagc acaacttgat ttaccaagac aagctgtggg ctgtatttgg 1200  
 tcctcaggct gtggcttgct agcccctgct ctaatggatt gtgttggcgg caatggcaag 1260  
 attcccttgc tgagaacatg cagatgaata accttatagc tgggatgcag agaggagaaa 1320  
 gttctgtagg taataagtgc actgtattcc ctctgagcta cttcctgggc atagttagat 1380  
 gtatggccag tgtgtgtact tgagccatca ggctcagaca cttgttgaat ataccttca 1440  
 ttgcattagt ttgtaattgt gctagtgtct ggtgctgagc tcagctgccc ggatgacctt 1500  
 tccagactga aggacatact cccctaacag ctgggagtgc tgctggctgg ccatttactt 1560  
 ccagccctta tgaggagtgt cccctgctga agagccctgc ctgccccaga tcataccccc 1620  
 ttcttgcttg taacccttac cggctccata tgggggtacaa agggctggcc tcctcacc 1680  
 aacttgggaa amcctctggg gccatcccag ctccagagcc ccttgtgggg tcagttagac 1740  
 ctcatgtgtg ccacattaca gccagtgcc tctccctgac aagcctgtac ccagccggct 1800  
 cagcccacag cactgtccta tgaaccttcc tgcacgccat tctccacctc agtatctgct 1860  
 ttcgggggaa ccaacctgcg acagtgtctt tgtgtgtttt cagtcctgca ggtttgaact 1920

ctgacttttg	agacttttcc	agttatctcg	tggaatgaca	gattgtgcct	gtatgatatc	1980
aggcacaatt	aatgcaaatt	aggcagacct	at ttatttcta	gatgtgaatt	tgctatttat	2040
tttaaatgtt	gatttttatgt	gttatgtgac	tctgaactta	ttgccaaata	aaagtttgaa	2100
ttgtaaaaaa	aaaaaaaaaa	aatntctcg	gtccgcaagg	gaattc		2146

&lt;210&gt; 36

&lt;211&gt; 1384

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 36

gagcgtggat	aatccagaaa	aggggaaaat	tgaaaattag	tagtgttggtg	tggaggaact	60
gacactgaat	tagtgtggtc	tttttatgca	ttcggccatt	gttttgtcat	tgctcctcaa	120
ttgtttccct	actgctggac	ggaaaattta	gattgtctta	ttcagaaaac	caaatgcctt	180
tctattgtct	ttccttatta	ttattatttt	tcaaattaag	ttgatgtctc	ttttgtcagg	240
cagttgaaaa	atatgtttta	tgaggattgt	gggttttgtt	agttcttacc	acactgccac	300
gccacactca	gtttgagaaa	tacacacacg	acaactccag	actcatttca	gaaatatttt	360
tatccatgtt	tacctctgca	gctggtgcag	atctcagggtg	tgcaagaaat	atttctttaa	420
aaaaaaaaaa	acaaaaaaca	aatgctgtt	ttatttgtat	tttargacat	ttctgcctaa	480
gtcatctggg	tartcagaa	atctctgttc	actgcctggg	ataggtttat	gcaattttaa	540
atgttacata	aatgaatgaa	ataaggtgaa	catagtcatt	ttttaaaaaat	agcattatta	600
tttttatgaa	aaataaatag	aatgcttttg	attcataaaa	aggctatatt	tgcaaagtac	660
ttaactgggt	atgacattgg	ggaaaaatgt	ttaacattga	tgataattct	gctctcagat	720
ttgaaactgc	cttcagattt	ttgttctgct	ttagaggaac	aaaaatggaa	actcgggtga	780
attacgatgt	tgtttgtgaa	aagacatgtc	tcaaaactct	agctaactctg	tccaaaaaaa	840
aaaaaaaaagc	aacagtcccc	atactaaaaa	taccaatgaa	acaaaaaagc	cccatttgat	900
cttaaacata	tatacattta	gaagttttaa	gttaaatatt	aaggttatgt	gtgcatttaa	960
aaaattatct	tactgattga	ctttaagaag	ttaaccacc	aactactgggt	tcttgtcttg	1020
acagggtctt	tatctacact	gacaaaatga	aatactatgc	aatcattaaa	cattctattt	1080
tattttactc	tattttctcc	tattttatga	tagcatttta	taatggtata	tgaaggtgct	1140
catttatatgt	ggttcagtga	gagagggtgg	ctatagaata	gaatgtactg	taccttccca	1200
tttttgtaaa	aatatgttta	gaacaaagac	gaatacgata	catgttaatg	gtgggtagat	1260
tatggctgat	ttttatgctc	tttattttta	taattgtact	attataaatg	ctttactttt	1320
gtaataaaaa	ttaataaaaag	tcattttttta	aagaaaaaaa	aaaaaaaaaa	aaaagggcgg	1380
ccgc						1384

&lt;210&gt; 37

&lt;211&gt; 739

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 37

cagcgcagaa	ggaccagcag	aaagatgccg	aggcggaagg	gctgagcggc	acgaccctgc	60
tgccgaagct	gattccctcc	ggtgcaggcc	gggagtggct	ggagcggcgc	cgcgcgacca	120
tccggccctg	gagcaccttc	gtggaccagc	agcgcttctc	acggccccgc	aacctgggag	180
agctgtgcca	gcgcctcgta	cgcaacgtgg	agtactacca	gagcaactat	gtgttcgtgt	240
tcctgggcct	catcctgtac	tgtgtgggtga	cgtcccctat	gttgctgggtg	gctctggctg	300
tctttttcgg	cgcctgttac	attctctatc	tgcgcacctt	ggagtccaag	cttgtgctct	360
ttggccgaga	ggtgagccca	gcgcacacgt	atgctctggc	tggaggcatc	tccttccccct	420
tcttctggct	ggctgggtgcg	ggctcggccg	tcttctgggt	gctgggagcc	accctgggtg	480
tcacggctc	ccacgctgcc	ttccaccaga	ttgaggctgt	ggacggggag	gagctgcaga	540
tggaaccctg	gtgaggtgtc	ttctgggacc	tgccggcctc	ccggggccagc	tgccccaccc	600
ctgcccctgc	ctgtcctgca	cggctctgct	gctcggggcc	acagcgccgt	cccatcacia	660

gcccggggag ggatccccgcc tttgaaaata aagctgttat ggggtgtcatt caaaaaaaaaa 720  
 aaaaaaaaaa aaaaaaaaaa 739

<210> 38  
 <211> 743  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (1)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (5)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (25)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (41)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (44)  
 <223> n equals a,t,g, or c

<400> 38  
 nattncgccc aaggctccta aattnccgac cccacctata nggnaaagct ggttaccccc 60  
 tgcaggtacc cgggtcccga aattccccgg gtcgacccca cgcgtccgaa atgggtctaac 120  
 atatattgagg acaatggtga tgatgctcca cagaatgcta agaaagctag gcttctacca 180  
 gaaggggagg agacggttga atcagatgat gaaaaagatg agcatacttc taaaaagcgc 240  
 aaagtagagc caggagaacc agcaaagaag aaaaagtaga aacaaatgac cagaatttct 300  
 gtactgctaa acttggttga atgttttctt ggacagatta agttgatatt gtgggttatt 360  
 atgccacatc tccatgaaaa tgcatacgtt aatgaactaa taagtattgc ctcaagaact 420  
 ttccactata gaattctttt tttattttaa acatgtatgt attttaaact caactggtga 480  
 cttgtgattg tgaaattgat aacacttgga tgcattcttg ctctcacaga attggtgaca 540  
 tgctttgaga gttttgtcac atgttgacat gccaatgttc ttataaacct ttataaagg 600  
 aatatatttt taaagtaa atattgta atgtgtgaa cttgtagggt gcttttcaac 660  
 agtctttgta cagtgtaa atgatcatgga aataaaatta cttattcaat attaaaaaaa 720  
 aaaaaaaaaa aaagggcggc cgc 743

<210> 39  
 <211> 1066  
 <212> DNA  
 <213> Homo sapiens

&lt;400&gt; 39

cgcgtccgat	tttttgtatt	ttttctaggg	acggttttcg	ccatggtggg	caggttgggc	60
ttgaactctt	gaactcctga	cctcagggtga	tccacccgctc	tcagcctccc	aaagtgtctgg	120
gattacagca	tgagccactg	gccgactttt	gcttccttta	cagattactt	tgcagattga	180
catccaggct	atacccttgt	ttatgtgagc	gaggctgtgt	tgtaggcact	ggcactatac	240
ttacacagaa	gttggcaact	tctcattgcc	ttgcacaaat	gtggataata	gtaataacaa	300
actaatgttg	aatgattaga	atthtaatca	gttcttttaa	agtagccatt	atthttggcat	360
gttaggtagg	gagagattca	aattatthtat	gttatcttcc	caacttgaga	atggcattat	420
tagatcaagc	cattagactt	atagattgcg	gaagatttag	atgtacttct	taattatgaa	480
aattacattt	accttactca	aataacccca	tatataccta	aaagthttaa	taaactaaag	540
catgaccaca	tgactcaatc	agctgtccgt	aataacctgt	ggtataggta	ggacttagtt	600
gggcaaagg	tgagatgtgg	ccatcttcac	gggtctataa	gagataacca	aggtcccttc	660
aggtcggcca	gcctcatcat	gcctcagctc	tttctaatac	tgcttattag	taagggtggat	720
gtttattcat	ttattttatt	ggttagttaa	ataacaggcc	gcaatttact	gtgttactta	780
aattthtact	aagtaaaatt	taagaagaag	aagaagatga	tgactgtaaa	aacaaagggtg	840
tgaatatatt	tagctataca	tgagatagga	tatcaaatac	ataaatagtt	caatttgtag	900
ttgcacagtg	ccatttcacc	caggttaaac	aataggaaga	ctcgatgcat	acgtgtcatc	960
tgtaaatgct	ctcatgtgaa	agaaattgtg	aagtaaaata	aagtatcttg	agtagaaatt	1020
tttctacatt	ttctcaaaat	aaaaaaaaaa	aaaaaaaaaa	aaaaaa		1066

&lt;210&gt; 40

&lt;211&gt; 1144

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 40

ggcacgaggt	tacctgcagc	agttggagag	tgaaatagat	gaactttaca	tccaatatat	60
caagcacaat	gatagcaaaa	atatcttcca	tgcagctcgt	acccagcca	cactgtttgt	120
agtcactctt	atcacatatg	tgattgctgg	tgtgactgga	ttcattgggt	tggacatcat	180
agctagccta	tgcaatatga	taatgggact	gacccttate	accctgtgca	cttgggcata	240
tatccggtac	tctggagaat	accgagagct	gggagctgta	atagaccagg	tggctgcagc	300
tctgtgggac	caggctttgt	acaagcttta	cagtgcagca	gcaaccaca	gacatctgta	360
tcatcaagct	ttccctacac	caaagtgcga	atctactgaa	caatcagaaa	agaaaaaaat	420
gtaatgcaaa	ttttaagaaa	tacagggtgca	tgaccaattg	tcaattaaat	attcagtttt	480
atgtctccat	gcaaacattc	aaagtgcctc	catcagaacg	gataaaatac	taaacacctc	540
tgaagactgc	aaactggatt	agttctttta	cttcagtgtt	taataagcag	atgtatgtat	600
gcatgggtat	actattttgt	taacatgtac	aatttcctga	tttttcttca	aaaatgctgt	660
tataaagtat	ttgtctatth	atgataacag	tacacgtgtt	ctgcttgaat	ttactaaatt	720
ctactactgg	gttataatta	aatcatgtga	tattccacgt	ttggatatgc	tcatttaatt	780
tctacagaaa	aaattthtaa	ttattthaca	ttagccattt	gttaaaacac	agcatcataa	840
ctcagcaggc	tggattthaat	ctgtatcatc	ttatatatat	cacaatctta	ttthtaagca	900
cattthtagag	ttccttagtt	gctthtatcaa	aaaccagata	ttgctthttac	atggththaat	960
agaatataaa	cctcttgata	aaaaatgcac	aaaaaatcac	tttgtatatg	tgagthttcac	1020
tgcattgtat	atthththcat	ttggtacaca	aagaatgtat	tcttcatagg	thtattctth	1080
taatatgtga	actattatta	aagththactc	tggtthcttaa	gattaaaaaa	aaaaaaaaaa	1140
aaaa						1144

&lt;210&gt; 41

&lt;211&gt; 1187

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 41



gagatgggtc	tcctggtatt	agaaaggaga	gcagacgcct	agagtctctga	cctcttgggc	60
ctcccctaaa	caaaaagcat	agccccttca	ctaccctagt	aacaagtatg	acctgttgcc	120
tgttgaggtc	aggtgtacct	gtgtgtctct	ctctgtgatg	tatgtgggct	tggtgtccac	180
atcaatgtgc	gtgaatactc	ctaccacacag	actgctactg	tgccaatata	gaggytcata	240
agaaagataa	aagagaacca	gcagaacccc	aggttctgcc	tctccttcct	ccctttgccc	300
agataaaggt	gtacagtgtt	tgtggtgagg	ctgctggaca	ctgctgtttt	gggctttcta	360
gagagaatgc	aaacatgcaa	aaacagtgtg	gtagtgtagg	cagaacaaga	tctagaatca	420
gaagacctgg	agtcgaatct	cagctccaac	acccactgga	tgaacgtgga	caagctgctt	480
gatctctcag	atcctcgatt	tcttcatctg	taagatggaa	acaatatgtg	tcttttctaa	540
ggggaaagat	caaatgacat	catgtatgtg	aaaatgctta	gaaatgggtg	atgatgatga	600
tgatgatact	gattattagg	gtatctaata	aagttgggga	aagaacaaaa	tccccccctc	660
acaccctatt	gcaccaccag	ccttctccct	ccctgctcaa	ttatacccat	gttataggga	720
aggaaatagg	gaagaaactt	gtctaagggtc	acatggatgt	tcgcaacaat	tgggatgcag	780
aacttgatct	gactcccagg	ccaggtttga	ttttgatttg	cagtggcaga	gaggactaat	840
tttttactag	gttctgagcc	ctgtactgaa	ctctggattg	ccctgtactt	ggtggtggta	900
tacagttcag	cagtttttag	catatacacg	aatttggtgca	gccayaacca	ctatctaatt	960
ctagaacatt	tttatcacc	ccaaaagaaa	tcatgtaccc	atttgcagtc	acttgccatt	1020
ccctcttctc	cccagccctg	ggaaaccact	gacctacctt	atatctctat	gaatttgcct	1080
aatctggaca	tttcatataa	atgcaattgt	acatgtgaaa	aaaaaaaaaa	aaaaaactcg	1140
agggggggcc	cgagtaccca	attggcccta	cgaagaggcg	aacagag		1187

&lt;210&gt; 42

&lt;211&gt; 1220

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (958)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 42

aaccaatcaa	ctatcaagca	gtagtaatct	atcttttaaat	ttctactcgg	tcatgcaagt	60
aaaaatgcca	tcacgccttc	aatcccaaaa	ttcccaaatt	cccaaagtga	aaaaagtcca	120
tgcaacaaaa	tgagaaataa	gaaccgttac	aactgagatt	actggagtct	cttccggaag	180
acaaattatc	actgtaatat	gtaatatgaa	catgtgaagg	tctggctcat	gttttatgat	240
acacaggctt	caaaagttgt	atgagctcag	atagagactg	ttgtttgcaa	tttgttcttt	300
cactaattag	ctatgtgatt	tagttaaaaa	gttcaaaaga	ttgtagtaac	gattaaacaa	360
atgtgtggaa	gtttgacctt	gcagtgtctg	acccactcat	gatatatatc	caaccaatac	420
tagtttagttc	tgattacctt	aaaactggct	gctcaactta	caaatcttag	gtcaaatcac	480
aggccctcta	acgactacat	ataagaaaaa	cctatcacca	acgatcataa	ggtgccatat	540
caccaaggat	cttctatgcc	atgcagaccc	aatcaatata	aggcagaatt	gatttaatac	600
agtgtttctc	taagtatgaa	taggcaggta	tcctaatacat	taagaaaagg	tgctttggaa	660
gcagaaaggc	tgcgttcaaa	ttctggctct	atcacttacc	atctacctaa	atttaggcag	720
ttcttttatt	ctctgagcct	acttaactta	tctgtattat	tcctataccc	cactgagtta	780
ttgtgaggat	taaataagta	cagagaaagt	gtgtggaaaa	gtgcttcaat	aagaatgtta	840
gctattgcca	ctgctattat	tatgctccct	aaagaatatt	ccatgaccca	ataaacagga	900
aacactgcta	gagagtcaca	tgtgtttattg	catttgacat	cctgaggaaa	aaaaaaancc	960
ctattttaatc	aaataattct	taaattttatt	taactctggc	atccttttct	caaataacat	1020
taactactgt	ggtagtaatg	tgccgaagaa	tactcttaag	tgcaatgaac	actataaatt	1080
ttctataaag	ctgtatatct	aaacatacca	ctgtaccaca	tagcctatct	ctagtccatt	1140
ccagatgcaa	tcaaaatttc	cctggcctca	aaaattagaa	gactaaaact	ttcattacct	1200
cttgttcaga	aggattggtt					1220



<210> 43  
 <211> 1285  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (315)  
 <223> n equals a,t,g, or c

<400> 43

cttgatacat	ttgttgaaaa	tgaattgatc	atatatgtga	gtgtttatatt	ctacattctc	60
tattatattc	cattggtcta	tatgtccatc	cttgtgtcct	agagaccaca	ttgtgtcagt	120
accacattgt	tttgattact	gcagctttgt	aaaagttttg	aaaatggcag	gcctgagtct	180
tccaactttg	ttctttttca	atatgatttt	agattgtttg	gggccctggg	aaaattccat	240
ttacatttta	aaatcagttt	ttacatttct	gcaaataaaa	aggcccttgg	gattttgata	300
agggttgcac	tgaanctgta	gatcactggg	taatatcgcc	aacttaacaa	tatttaaatct	360
tccaatctgt	gaacacagga	tgtatagtcc	tatttattta	agtcttcttt	aatttttgtc	420
aacagtgttt	tgtagtcttt	agtgacagaag	tctttcacct	tcttgtttta	atttattcct	480
agctatttta	ctctttttga	ygctatygta	agtggaattt	taaaaattga	cttccttctt	540
gggctgttca	ttcctggtat	atagaggcac	aaatgatttt	cttaaaataa	atgttgaaca	600
tgtaccctgc	aattttgctg	gacttggtta	ttagctctac	tagtttgtgg	attccttgat	660
tttctagata	gaagatgtca	ttattgaata	aagagtttta	ctgcttcctt	tccaatttgg	720
atgctattca	tttacctggt	tgtttattta	tctgcctaata	tgtcatgggt	aaaactttca	780
gtacattggt	caatagcact	gggtaaagca	tacagccata	tcttactctt	tgttatggga	840
gaaatttttc	agtttttttt	ctcatthaagt	atattagttg	tggatttttc	acaaatcctc	900
tttaacacat	tttgaaagt	ccatgttttt	atcctgaaag	gttgtagaat	tttgtcaaat	960
gcttgttctg	agtgaattga	aatgattgca	tggatttttt	atccttcatt	gtattaatgt	1020
ggtgtattac	attgattgat	actcttgkac	tgaaccaccc	ttgkattctt	gagataaatt	1080
ccatgtggtc	atgatataata	atccttttta	atatgctgct	gaatttggct	tactagtact	1140
ttgktaaata	aaagatttta	gtatctagtt	tgtcttttct	tgtgggtatat	ttgtctggct	1200
ttggaatcca	tgtaatactg	gcctcataga	atgtattgca	aagttttttc	tcctctgttt	1260
tttgaagag	tttgagaggg	attgg				1285

<210> 44  
 <211> 985  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (372)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (386)  
 <223> n equals a,t,g, or c

<220>  
 <221> SITE  
 <222> (391)  
 <223> n equals a,t,g, or c

&lt;400&gt; 44

aaccaatcct	aatccctcct	aacccccactc	atagtccaga	gaagataaaa	agttcaggga	60
catggtgtac	tgctctccag	atcctttccc	agctgaaact	ctatacccat	ttagcactaa	120
tttcccattt	tccctagccc	aagataacca	accaccattc	tactttctct	ttgaatttga	180
ctcctaggta	cctcacatat	atattcctat	ttagtggaat	tgtagactat	ttatcctttt	240
gtgactggct	tatttcactt	agcactgttt	tctaggttca	tccatgttgt	agcttctgtc	300
ataatttcct	tcccttttaa	ggctgaataa	ttttccatgt	gtatatgtac	catatttggt	360
catccattca	tncatgatag	atactngagt	ngctctacct	tttggcaatt	gtgagtgatg	420
ctggttatga	acatgggtgt	acaaatgtct	gttttgagtc	ctkgtttkgc	tttcgtgcat	480
tctattaaaa	atgttaaggt	tcttkgcttt	ttactttttt	ttctctttta	tccacctttt	540
gactttattc	ttaaagtttw	atcttggtat	aataatttat	ggatgttgtc	agaattagga	600
gcaaagtccc	gtgtgtcctt	cactcagttt	cccccaatgg	taacattttt	cataattaca	660
gcacaatgtc	aagaccggga	aattgacatt	gatacagtc	atatacaggt	gtcccttggt	720
atctgtgggg	gattgggtcc	agaaccctcg	atgataccaa	aatccatgga	tgcttgagtt	780
tattagataa	aatggcatag	tgttttcaca	taacctacac	acatcttcct	gtatacttta	840
aatcatcttt	cgattactta	taacacctaa	tataatataa	atgccgtgta	aatagttata	900
ttgtattgat	tttatttgta	ttttttttat	tgttgttatt	tttcattgtg	ttttttccct	960
gaatattttc	tatctgtgat	tggtt				985

&lt;210&gt; 45

&lt;211&gt; 2610

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 45

gcctgaagga	ctgcctcggt	tcaacaacaa	ctttatggct	cccggaaagt	cctcctcccc	60
gtcccccttc	tttccagcct	cacgcccgtg	ggctgcagtt	ggaacgatgg	cggcggcagc	120
tgccgcccgg	cctagcccgg	ggctctggacc	tggggactcc	ccagaagggc	ccgaggggga	180
ggctccggag	cgtcggcgga	aggcgacagg	gatgctgaag	ctttactacg	gcctctcgga	240
aggggaggcg	gcgggacgcc	ccgcggggcc	cgaccccctg	gacccgactg	atctgaacgg	300
ggcgcacttc	gacccggaag	tttacctaga	caagctgcgt	agagagtgcc	ctctggccca	360
gttgatggac	agtgagacgg	acatgggtgcg	gcagatccgg	gctctagaca	gcgacatgca	420
gaccctggtc	tatgagaact	acaacaagtt	catctcagcc	acagacacca	tccggaagat	480
gaagaacgat	ttccggaaga	tggaggatga	gatggaccgg	ctggccacca	acatggcagt	540
gatcaccgac	ttcagcgctc	gcacagcgcc	cacgctgcag	gaccgccacg	agcgcatcac	600
caagctggca	gggggtccacg	cgctgctgcg	gaagctgcag	ttcctctttg	agctgccctc	660
gcgcctcacc	aagtgcgtgg	aactgggccc	ctatgggcag	gcgggtgcgt	accagggccg	720
cgcgcaggcc	gtgctgcagc	agtaccaaca	cctgccctcg	ttccgcgcca	tccaggacga	780
ctgccaggtc	atcacggccc	gcctggccca	gcagctgcgg	cagcgcttta	gggagggcgg	840
ctcaggcgcc	ccggagcagg	cagagtgcgt	ggagctgctg	ctggccctgg	gcgagcctgc	900
ggaggagctg	tgcgaggagt	tctggcgcac	gcccgcggcc	ggctggagaa	ggagctgaga	960
aacctggagg	ccgagctggg	gccctcacct	ccggtcccg	acgtgttaga	gttcaccgac	1020
catggaggca	gtggcttcgt	gggcggcctc	tgccagggtg	cggcggccta	ccaggagctg	1080
tttgcgccc	agggcccagc	aggtgccgag	aagctggcgg	ccttcgccc	gcagctgggc	1140
arccgctatt	ttgcgctggg	ggagcggcgg	ctggcgaggg	agcaggggtg	tggtgacaac	1200
tactgctgg	tgcgggcgct	ggaccgyttc	caccggcgct	tgcgggctcc	cggggccctg	1260
ctggccgctg	ccgggctcgc	agacgctgcc	acggagatcg	tggaaacgagt	ggcccgcgag	1320
cgctggggcc	accacctgca	gggtctccgg	gcggccttcc	tgggctgcct	gacagacgtc	1380
cgccaggcgc	tggcagcacc	tcgcgtggct	gggaaggagg	gccctggcct	ggccgagttg	1440
ctggccaatg	tggccagctc	catcctgagc	cacattaagg	cctctctggc	agcagtgcac	1500
cttttcaccg	ccaaagaggt	gtccttctcc	aacaagccct	acttccgggg	tgagttctgc	1560
agtcagggtg	tccgtgaggg	cctcatcggt	ggcttcgtcc	actctatgtg	ccagacggct	1620
cagagcttct	gcgacagccc	tggggagaag	gggggtgcca	caccacctgc	cctgctcctg	1680

ctgctctccc	gcctctgcct	ggactacgag	acggccacca	tctcctacat	cctcactctc	1740
actgatgaac	agtttctggt	gcaggatcag	ttcccagtga	cgcccgtgag	cacgctgtgt	1800
gcagaggcca	gggaaacggc	gcggcggtg	ctgaccact	acgtgaaggt	gcagggcctg	1860
gtcatatcac	agatgctgcg	caagagcgtg	gagactcgcg	actggctcag	cactctggag	1920
ccccggaatg	tgcgggccgt	catgaagcgg	gtggtggagg	ataccaccgc	catcgacgtg	1980
caggtggggc	tcctgtacga	agaggggtgt	cgcaaggccc	agagcagcga	ctccagcaag	2040
aggactttct	ccgtgtacag	cagctctcgg	cagcagggcc	gctacgcccc	cagctatacc	2100
cccagtggcc	cgatggacac	caacctcttg	agcaatatcc	agaagctatt	ctctgaacgt	2160
attgatgtgt	tcagccctgt	ggagttcaac	aaggtgtcgg	tgctgaccgg	catcatcaag	2220
atcagcctga	agacgctgct	ggagtgtgtg	cggctgcgca	cctttgggcg	cttcgggctg	2280
cagcaggtgc	aagtggactg	ccactttctg	cagctctacc	tgtggcggtt	tgtggccgac	2340
gaagaactcg	tgcaactgct	gctggacgaa	gtggtggcct	ctgctgccct	gcgctgcca	2400
gacctgtgct	ccatggagcc	cagtgtggtt	gaggtcatct	gcgagcgcg	ctaggcgag	2460
ccgctgccat	gcaccggtct	gtccctgcac	cccatggcac	ccaggatctg	gtctcggtgg	2520
tccttccccg	caggcaggtg	tcaggaccgg	cctaataaac	atgtgtggcc	tcctcaaaaa	2580
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa				2610

&lt;210&gt; 46

&lt;211&gt; 951

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 46

ttctatagca	ttgcatttct	taagtaacta	catacacaca	taaatatatc	actatcaaaa	60
aacatttctg	tcacctgagg	gagtttcgaa	ttttctatgt	ctcttcctag	tctctttccc	120
caaacctcta	ctaaccatag	ttaattttgt	ctgaggattt	tatataagtg	aaattataca	180
gtaagtgtaa	agtttttgag	atctattctt	gatgtacttt	atctgtaatt	tgttgctttt	240
acttggacat	atatatgtgt	atatatggat	gtctgtctat	atatatatgt	ataagtaaaa	300
acttttagaaa	ttttagaaag	aaagctttta	gaaatagagc	tactatgatc	actctcgaac	360
caggtagtct	ttctttgttc	atatgctttc	ttttctctta	agaaagtatt	aaggagtgga	420
attgctgagt	catagcaaaa	aaaatgtatg	tttaacattt	taagtagctt	ctaaatcatt	480
tttcaaagtg	attgtgtcat	ttcacactcc	cctcatcagt	atatgagagt	tttcacttaa	540
tttcagcat	ttaagattgt	cagtctttca	ttttaatcat	cttcttttga	tgcttctttt	600
tgtctctgtt	ctctaattaa	tgaaaaacat	tgggatcaca	tcaagagata	taaaaaaaaat	660
tatgcctttc	caaggattaa	cctttgtgaa	atttcagtgt	gtcttgtttt	ctcactataa	720
atcaacaggt	ctgctgcacg	agagtcagaa	catcactaca	aatgaaaaa	cctgcaccag	780
gagagaagcg	agaaggcact	gctaagggac	actcgattta	ctcaagatgg	aataagtaca	840
aagttacgta	gtcaggagaa	ctagtacgaa	aaaggaatca	ggtctgtgtc	agagtaaaaa	900
ttgcatctag	caaattttaa	caggcaaaaa	agagttattc	aaggccaatg	c	951

&lt;210&gt; 47

&lt;211&gt; 554

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (16)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (26)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (94)

<223> n equals a,t,g, or c

<400> 47

aggtcactca	ttaggnaccc	ccaggnttta	cattttatgc	ttccggctgg	aatgggtgtgg	60
gaattgtgag	cggataacaa	tttcacacag	gaancagcta	tgaccatgat	tacgccaaag	120
ctcgaaatta	accctcacta	aagggaacaa	aagctggagc	tccaccgcgg	tggcggccag	180
ctctagaact	agtggatccc	ccgggctgca	ggaattcggc	acgagctact	aataattatt	240
acattttgac	taccacaact	caacggctac	atagaaaaat	ccaccctta	cgagtgcggc	300
ttcgacccta	tatccccgc	ccgcgtccct	ttctccataa	aattcttctt	agtagctatt	360
accttcttat	tatttgatct	agaaattgcc	ctccttttac	ccctaccatg	agccctacaa	420
acaactaacc	tgccactaat	agttatgtca	tcctctttat	taatcatcat	cctagcccta	480
agtctggcct	ataaaaaaaaa	aaaaaaaaaa	ctcgaggggg	ggcccggkac	ccattsgcca	540
aakggggggg	ttta					554

<210> 48

<211> 3052

<212> DNA

<213> Homo sapiens

<220>

<221> SITE

<222> (2994)

<223> n equals a,t,g, or c

<400> 48

agcgctcccg	cccaccctcc	ggctctcggcg	gctctccaga	gcgtctgtaa	acacccagag	60
actgtcatgg	agggggagga	ggaggcgggc	gcggcgaagg	gaggcgtttg	gggccgcctc	120
cagggctcgc	tctgccattc	ctgaactggg	ccctcgtccc	cgtgactctg	gcatcagggg	180
agcgaactgt	taggcgagag	gaggaggcag	ccagaaccat	atcccccttct	tcctcggggc	240
gggggcccgg	ccaggccggc	tgagccgggg	gagggtcccg	ggaggggagt	cctggccagg	300
ccggcctgtc	tgccgcgatg	gatgacagta	aggtgggttg	aggcaaagta	aagaagcccc	360
gtaaactgtg	tcggaagcca	gccccaaattg	acttgaaagc	aaaacttgag	aggagccggc	420
agagtgcag	agaatgccga	gccccgaaaa	agctgagata	tcagtatttg	gaagagttgg	480
tatccagtcg	agaaagagct	atatgtgccc	tcagagagga	actggaaatg	tacaagcagt	540
ggtgcatggc	aatggaccaa	ggaaaaaatcc	cttctgaaat	aaaggcccta	ctcactggag	600
aagagcagaa	caaatctcag	cagaactcaa	gcaggcatat	caaggctggg	aagacagatg	660
ctaatagcaa	ttcctgggtga	agattatata	aagatgagtc	agtgattgaa	gccaatattc	720
tgattcccat	ggaagatgga	tgggcaagag	tgtacttctt	ggctccattt	actacctact	780
gctcagtagt	catctctgta	aatctgcaat	ttctaccaa	atgtgtgate	gtagatctca	840
aaggatcttg	ctttaacttt	caacacttag	aaaatctaca	aacattcaga	cctgtctggg	900
ttgggtattgc	cacccatgac	atttaacatg	ttgtgatgct	tgaaaacaca	ggagtagaga	960
aaatcgatga	agattgtatt	tttgcacctt	aactccacat	tgctttattg	gttaatttat	1020
attctttcca	tgtaattcat	gtaattgtat	gtctgtgtgt	gttttatgtg	tcaccacctt	1080
tcatgttttt	gattgcccta	caaagagaaa	ccaaatgagc	tgattactga	ctataagttc	1140
tcagccttta	tggacctaat	cttattttta	tttacttgag	taatgtttat	tctctgcatg	1200
aaccatgatt	tctcctgtga	gccattccag	cataagctgt	gaatatgtat	taacaaatat	1260
atacatttct	atttttataa	tccataagga	tatgcctgtt	ttaaataaca	tacatattaa	1320
caatatctat	caggaaaacc	ctcaagacag	cttctagtta	aaacctttgt	tgctgtcctc	1380
tcaaactata	tttataaaaa	tttgctaggg	ccaaatccat	acttgcagaa	taattcatca	1440

aattttat	tttaagt	gaaa	agtaac	cttt	caggcatt	tc	agcagcatac	attgacaatc	1500
tagggatat	atgtat	gtat	gtttct	tatt	gtatgtct	at	atatgtatgt	ggggaggaca	1560
ggagtgaat	gttcacac	act	tttctt	gcgt	actcaact	aa	attggagaat	gtttctgaag	1620
aaaattggat	gaaattag	ct	gctgag	attg	agtttct	gcc	ttaaaatctg	aaacaaaaaa	1680
agggacaaat	tgctggtaga	tctact	gact	gtagcc	atca	ccagaacact	tagtttcttc		1740
ccagacatga	atttcctgac	aggctctgag	ccagaaacac	actgtggg	cg	tgcatctggt			1800
tcagccctgg	atatgcctcc	actgtggagg	gaagcattgg	aactggag	tt	gtgcacagca			1860
ccgcagctag	tgggcctggc	tgggggtgca	gaagggcagg	gtagttct	ttt	ccatttcccc			1920
ttgcagctac	acttactttc	tccctgacta	ccaaatggag	tgacgg	ttga	agaatcaagc			1980
tcataatatg	atacacccag	acagacccag	aaatcttttg	atttccccag	cgagatactt				2040
accactctct	cttctctttc	ccatgggttaa	aacaggag	tt	gtgttctctg	tcgctgctaa			2100
attattcttc	acttgagtga	agcaaata	tatt	tagcttccaa	ataaagtata	ttttgttctg			2160
aactggcagt	tttaagaggag	ataccgtagt	ctagattgtc	tttgaggtag	aatatatgat				2220
ggaccacata	gtttaagatt	acaagg	tact	ttccctcttc	tctcaggcca	tgcatattgaa			2280
ttaaagtgcct	cccttagcct	gagctcctaa	agacctcttg	tacttgggtg	caaactcaga				2340
tctctagtct	agtctagtct	agtctagtca	tgtgtgactg	tgtactgaga	taccactagc				2400
ccagtttgct	gggtctctat	gttttggaag	ttgggggttt	aataatattt	aatgtctttt				2460
taggatcata	aacataaata	ggtttggtga	tattcaggcc	ttcagaattt	aagggtttta				2520
aatatagcat	tcagattgta	attgggtatgt	ttttgccatc	tggtccactc	aattgtatat				2580
ttttattttg	caaaacaacc	caaaacttac	tttatgttgc	tttgttcagt	accttttgaa				2640
ttcccccaga	agagttgttt	tcaaaacaaa	cattccagaa	tgaatgtcgc	tcttcaatgg				2700
aacgtcttta	ttgtgcttga	agtatttctt	ctctgggtgt	aattctaaat	tacagagtca				2760
tttgagcttt	ctcccctccc	gctagtatct	ttacaggagc	agggaagagc	agtagaggat				2820
gtataatttt	gggcgaagt	aaattacaat	ttatttgagg	ttattcctaa	acctatttat				2880
ttgggtgttt	ggaggagatc	acacactaag	agaacgttga	ttgcctcggc	tattgtgctg				2940
gctggacact	ttggtcactt	ttgaagcatg	ttaataaatg	tcactgatta	aaanaaaaaa				3000
aaaaaaaact	cgaggggggg	cccaggtacc	caattggccc	tagaagaggc	ga				3052

&lt;210&gt; 49

&lt;211&gt; 2154

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (461)

&lt;223&gt; n equals a,t,g, or c

&lt;400&gt; 49

cccacgcgtc	cgarccgagc	tgaggctcgg	cttcctgctg	atggtcaggg	ttttggcaac	60
tccccggtgt	gagaggggta	gggagtgtc	ccggcgccga	cggggccgag	ttcaccagcc	120
gccggggcag	tagtcgaagg	cccggcgcg	catgtcctgg	gtgccgcggt	gcgggcagtg	180
aacgcgcgcc	gggcgggatg	ggccggcgcc	gggcgccaga	gctgtaccgg	gctccgttcc	240
cgttgtacgc	gcttcagggtc	gacccacaga	ctgggctgct	catcgctgcg	ggcgaggagg	300
gcgccgccaa	gacaggcata	aagaatggcg	tgcactttct	gcagctagag	ctgattaatg	360
ggcgcttgag	tgccctccttg	ctgcactccc	atgacacaga	gacacgggcc	accatgaact	420
tggcactggc	tggtgacatc	cttgctgcag	ggcaggatgc	ncactgtcag	ctcctgcgct	480
tccaggcaca	tcaacagcag	ggcaacaagg	cagagaaggc	cggttccaa	gagcaggggc	540
ctcgacaaag	gaagggagca	gccccagcag	agaagaaatg	tggagcggaa	acccagcacg	600
aggggctaga	actcagggtta	gagaatttsc	aggcggtgca	gacagacttt	agctccgatc	660
cactgcagaa	agttgtgtgc	ttcaaccacg	ataataccct	gcttgccact	ggaggaacag	720
atggctacgt	ccgtgtctgg	aaggtgccca	gcctggagaa	ggttctggag	ttcaaagccc	780
acgaaggag	attgaagacc	tggcttttag	gcctgatggc	aagttggtaa	ccgtgggccg	840
ggaccttaag	gcctctgtgt	ggcagaagga	tcagctggtg	acacagctgc	actggcaaga	900

aaatggaccc	acctttttcca	gcacacctta	ccgctaccag	gcctgcaggt	ttgggcaggt	960
tccagaccag	cctgctggcc	tgcgactctt	cacagtgcaa	attccccaca	agcgcctgcg	1020
ccagccccct	ccctgctacc	tcacagcctg	ggatggctcc	aacttcttgc	cccttcggac	1080
caagtcctgt	ggccatgaag	tcgtctcctg	cctcgatgtc	agtgaatccg	gcaccttcct	1140
aggcctgggc	acagtcactg	gctctgttgc	catctacata	gctttctctc	tccagtgcct	1200
ctactacgtg	agggaggccc	atggcattgt	ggtgacggat	gtggcctttc	tacctgagaa	1260
gggtcgtggt	ccagagctcc	ttgggtccca	tgaactgcc	ctgttctctg	tggctgtgga	1320
cagtcgttgc	cagctgcctc	tgttgccctc	acggcggagt	gttcctgtgt	ggctcctgct	1380
cctgctgtgt	gtcgggctta	ttattgtgac	catcctgctg	ctccagagtg	cctttccagg	1440
tttccttttag	cttccctgct	tcctgggaat	caggagcctg	gacactgcca	tctctagagc	1500
agagtggagg	cctggactcc	ctttgctcac	tccattcggg	tccacagctg	aggttgcctc	1560
tgacaagatg	aatgggcact	gcctgccctt	ctagtgaata	ggcttggcta	tggccctgtg	1620
tgactccagg	tcccaggaac	cttgccttcg	tcactctgtg	atccatccag	aacagcggta	1680
tctgaagccc	aggccatact	ccctgcctcc	tttcttctgc	ctaccagagg	ctccagagtt	1740
gagcttgctc	ttatctagaa	acatgtgaag	atgcccaga	gcctggaggc	actgctgtcc	1800
ttcctgcaga	aacagtttct	cctcctcccc	tcagccttgt	ggccagttcc	tcttcacatg	1860
aagcccctgg	catttgctgg	ggaagggact	ggcctggtac	ttgctgttag	ggcaggaagg	1920
ggcaaaagga	agacttgggt	agtaatctgg	gggttcagat	gggtagcact	aagccagctg	1980
gcctaaagat	gcaataagtt	cctaggtagt	ctacccttac	cttgaggaat	gggaaaatga	2040
acctcagccc	attaggcagg	aaaagttgat	atttaataaa	caaggaaaga	gtgaactgag	2100
acccccaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaa	2154

&lt;210&gt; 50

&lt;211&gt; 643

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 50

aattccccgg	ggctctgagg	gccctccaga	cctgctcggg	tgctggggcc	atgccgagtc	60
gcggccctgc	tcagccggaa	gaggctcccc	gacctggatg	tacagggcag	tctctcttcc	120
cggggctatg	ggctgggcct	gtcctgcccgt	catggccccc	tgcttcctgc	tccttgagac	180
tggtcccg	actttgcca	ccatccatgc	agtggctccc	agggcagagc	ctctccttgt	240
actttggcag	ccatagaaag	cgtgctcatt	ttctgttttc	ctgtgttagg	aaaaaaccac	300
ctgttttcca	aggggagagg	gcggggcctg	agggtggggg	cggggcctct	tcattggccc	360
agcttggcga	aagcgaggca	cactgcttac	tgccttgggg	ttgtggagat	ggaccctga	420
cctcgtggag	gccgtgtggg	ggcagcagcc	tggcctgtgc	catggtgggt	gtcctggggc	480
ctgtgcggag	ggagccacct	caccctgcag	cccaggttgc	aggtgtggcc	ttgtttctcc	540
ttgcccagca	gtgctgcctt	cagcggccgt	gacggggcca	gctggacaca	cggtagagatt	600
ttctcgtatg	taaataaaaag	gcaatttggt	aaaaaaaaaa	aaa		643

&lt;210&gt; 51

&lt;211&gt; 682

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (665)

&lt;223&gt; n equals a,t,g, or c

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (673)

<223> n equals a,t,g, or c

<220>

<221> SITE

<222> (677)

<223> n equals a,t,g, or c

<400> 51

gagtcaatat	tttttggtt	tgtagataaa	gacttaaatt	tttgtgcaac	atagtgggtg	60
taaagcacac	atccattgga	ctatgcaa	caatttctac	tagcagcact	gagcaactgg	120
gcacgtgcta	gtacgcaccc	cctcctactc	taactgtgct	agctcttggg	ttgaggaatg	180
gaatgggtag	ccttctgctt	actggagatg	cacagagaaa	catggctgaa	aacagcttta	240
aaaagaacat	gtctttgggt	tcatgtgtgg	tgctgctgcc	cgggtgtaaa	ctatccccac	300
ccagggccag	ccattatcac	cagacctctt	ctgggcctgg	ctgtgaagcc	ctgttttttg	360
tattcagaat	ggatctttaa	ggctcagatg	ggctcaaatt	gaaaccttgt	gttttataaa	420
atggatgttt	agtaaaggaa	ctggttctca	gttatgttta	cagcacttgg	aattgtgtgt	480
tcttgtacat	tttgtatttt	aaaacctttt	atgggagtgc	agtgtcctt	acacaaatac	540
aaagggaga	agagccagca	gttgaggctc	ctcagtttta	gtgctgaaat	aataaacagt	600
gacaggtcaa	cagtaaaaaa	aaaaaaaaaa	aaamctcgag	ggggggggcc	cgggtacccc	660
aaatncgcc	tanaagnag	gc				682

<210> 52

<211> 4024

<212> DNA

<213> Homo sapiens

<400> 52

ccacgcgtcc	gaaagcctct	tgcccatcaa	tggtttgctc	tctggagaca	aagtcaagcc	60
agtcctcatg	aactcaaagc	tgctcttga	tgtcctgggc	agggctctggg	acctcagtga	120
cattgacaag	gatgggcact	tggatcgaga	tgagttcgct	gtggccatgc	acttgggtga	180
ccgagccctg	gagaaggagc	ccgtgccctc	cgccctgccc	ccgtccctca	tcccaccctc	240
caagagaaa	aagactgtgt	tccctggcgc	cgtccccgtc	ctgcctgcca	gccccccacc	300
aaaagacagc	ctccgctcca	cgccgtccca	cggcagcgtc	agcagcctca	acagcacagg	360
gagcctgtcc	ccaagcaca	gcctcaagca	aacacagcca	acagtgaact	gggtgggtgcc	420
cgtgcagaca	agatgcgatt	tgatgagata	ttcctgaaga	ccgacctgga	cctggatggc	480
tacgtgagtg	ccaggagggtg	aaggagatct	tcatgcactc	gggcctcacc	cagaaccttc	540
tagcacacat	atgggccctg	gccgatacga	ggcaaacggg	gaagttaagc	aaagaccaat	600
tcgcgttagc	tatgtatttc	attcagcaga	aggtcagtaa	aggcatcgac	cctcctcaag	660
tcctctcgcc	ggacatggtc	ccgccttcg	agagaggcac	gcccggcccc	gacagttcag	720
gctctctcgg	ctccggggag	tttactggcg	tgaaggagct	tgatgacatc	agtcaagaga	780
ttgccagtt	acaaagagag	aaatattcac	tggaacaaga	cattcgagaa	aaggaagagg	840
caatcagaca	gaaaaccagc	gaggtgcagg	aattacaaaa	tgacctagac	cgggaaacaa	900
gcagtttgca	ggagctcgag	gctcagaaac	aggatgctca	agaccgcctg	gacgagatgg	960
accagcagaa	ggccaagctc	cgagacatgc	tgagcgacgt	ccggcagaag	tgccaggatg	1020
agactcagat	gatctcatca	ctgaaaacgc	aaatccaatc	tcaggaatct	gacttaaagt	1080
cccaggaaga	cgatctgaac	cgagccaagt	cggagctgaa	ccgattgcag	caggaggaaa	1140
cccagctgga	gcagagcatt	caggctgggc	gagtccagct	ggaaaccatc	atcaagtccc	1200
tgaagtcaac	gcaagacgaa	atcaaccagg	caaggagcaa	actttcccag	ctgcatgaaa	1260
gccgccagga	ggcccacagg	agcctggagc	agtatgacca	ggtgctcgat	ggagcccatg	1320
gtgccagcct	gaccgacctg	gccaacctga	gcgaaggcgt	ctccctggca	gagaggggca	1380
gttttgagc	catggatgat	cctttcaaaa	ataaagcctt	gttatttagc	aacaacacgc	1440
aagagttgca	tccggatcct	ttccagacag	aagaccctt	caaatctgac	ccatttaaag	1500
gagctgaccc	cttcaaaggc	gaccggttcc	agaatgaccc	ctttgcagaa	cagagacaac	1560
ttaacagatc	catttgagg	ggaccctttc	aaagaaagtg	accattccg	tggtcttgcc	1620



actgacgact	tcttcaagaa	acagacaaag	aatgacccat	ttacctcgga	tccattcacg	1680
aaaaaccctt	ccttaccttc	gaagctcgac	ccctttgaat	ccagtgatcc	cttttcatcc	1740
tccagtgtct	cctcaaaagg	atcagatccc	tttggaacct	tagatccctt	cggaagtggg	1800
tccttcaata	gtgctgaagg	ctttgccgac	ttcagccaga	tgtccaaggt	aaagcccctc	1860
cacggagccc	ccgcgcctct	gctagtgtct	ttgtgcctct	tgtcatgggt	tgggctgcca	1920
ggcgtaattg	ttcatgtcac	gtatgtatct	ccccggcacc	tttccaacac	aaggtcaggt	1980
ctggaaagca	tccatggctg	tgatccaatg	cactgcagtc	ccgtgggggt	agccctgacc	2040
cttcccagtg	gcataggtgc	cctgggctcc	cctggctccc	actgggtgtct	gacgaccatc	2100
aggtctcaga	cggtgaagtc	attgccatgg	ccgagtagaa	acttgagaag	gcgttgggca	2160
caggcgtctc	gagagggcca	tgggcagcag	gcctgcaggt	cgaggcgtcc	agggaaagtg	2220
ggccagcgag	tgttcagagt	cgcccatgtg	acagaactct	caaggggtgt	tgaaaagtga	2280
gggccggcct	ccttgaataa	atgcagacag	atccacaggt	gaggagtga	ctgttgaagg	2340
aggacagcag	cgaggctggg	taccagactg	gacccgttgg	tgtccggccc	tcaccacacg	2400
gtgacctgaa	ctcactgtgg	ctctggctcc	agacagctgt	agggcttctc	cagacgctgg	2460
cagtgtgggt	ggagctcatc	tttaaggtag	catagagaat	tcaacaggga	acatccagaa	2520
tggacaagtc	aaaaaacaac	aatagaagcc	ccttacgact	gaggtgcctt	ccggaagaaa	2580
ctgctggaaa	gttggaaatg	ggggctgtgg	tgagtgggaa	ggacaggggt	gcggcaccat	2640
tgtgcttcct	tgtgagttgt	atgtttccct	gtttgagtga	ggagcatttt	tctttggcag	2700
gaatgaagga	aagcagatga	ggtgtgtccc	gtccctgaca	cctgcgtcca	ctctgtccct	2760
tctcggcagc	ccccaggcag	gggcagatgg	gctctgtcct	cgggggctgc	aggggcctgc	2820
tttctcacg	agttccacct	gcagaggccc	aggcattatc	cacgctgcgg	gaagaggcag	2880
tagccacaag	cagaacacag	ccccctcct	ggagcccatg	gcccagggtg	taaggggtgg	2940
ccggtgtggg	gggtcccctg	gagcagctgg	gcccagtcct	tgtgagcgct	gctgtcagcc	3000
tgaaggtggg	aagggaggga	gccccgggag	tacctggaag	acgcgtgccc	caggcatgag	3060
gccccctcgg	ggccccaccg	gccccatgcc	caggtgcccc	aggtgccagt	gtgactgggc	3120
agtctccctt	ggctgccccg	tgttcaactca	gctggtggag	cagatgagtc	tctgagcatc	3180
ttgggggttg	cagcggctgc	aacttgtgat	cacagaacac	agattccgag	aagcacttga	3240
gattctgttc	aggttttcaa	agacgcgtgt	gtgcgtgccc	tggcggcgca	gcagatccca	3300
cttctgaccg	aggcagggtg	tgtgtttaca	cagatgagcc	tgggcagagg	gagcccagag	3360
gccctggcat	ggcgcccccc	tccttccagc	acacggagaa	ggcctccatc	cctcatatgc	3420
cacgctgtga	cacggttttg	gactggcaga	agatttaggg	cccagtgcac	aggctcctgg	3480
gcaggcggga	aaccatctcc	cgaaagcaga	aaacgttggc	acacgggtgg	agcagtcctt	3540
caggcgtcct	ggtcgtttct	ggttgtcgca	agtggactcc	tcttcctttg	ggaggctgga	3600
tctgtcttcc	tcctctcact	gtgagcagag	ggagggagag	ctagcagagc	ctgagctagg	3660
gccgggatct	aatgaccgcc	ctctgcctgt	ccttgggccc	tgaaccagg	ggctcttcac	3720
acacatgcag	ctgtgatctc	agagccggtc	tcacttccta	ttggcgggtg	cgggccagga	3780
cagagttgga	cagagctgtg	gaaaatgggt	ctccaatact	atcttgtttt	gattccagcg	3840
ttccatcagc	tgccttcacg	acacttgtac	agccgcattc	tgagctcaac	cctgaccttg	3900
cagctctgtt	tttctttttt	ggttacctat	ctgtttttca	agaacaacat	tgtggatata	3960
aataaaatgc	taagggaata	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	4020
aaaa						4024

&lt;210&gt; 53

&lt;211&gt; 563

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 53

ctacatagcg	atagtatggg	ttgataatac	agttttttctc	ctctagctgc	ctcctgttct	60
tgaggtaatg	ctcagtgatt	aactagatag	gactagataa	tttatttcac	cttggcaggt	120
cttgcccca	agttggccta	ttcctatggg	atttcctttt	ctttcttttg	cttcttcctg	180
cacactaagt	atggatattt	ctgacagtat	ttttcttaat	tctttattta	tttttcttgg	240
taattttctg	ctgtagagct	ggtacccatt	agacataatg	actttccttg	acttgcttct	300
tgctcataaa	tttctgtaag	acaatcagga	agttttttaga	aacaaacagt	ggaaccaatt	360



aggctgttaa	gtgagacacg	cctaaattaa	gctatgagga	aataaatata	gctccagcaa	420
aatagaaat	aattttcaga	ttatccatgc	ttttgtttcc	ctggttttct	tcgaacctca	480
tttcctcaag	ctatgaagtg	aaatgaagtg	gtaagtacac	cgctatgtag	gaattcgata	540
tcaagcttat	cgataccgtc	gac				563

&lt;210&gt; 54

&lt;211&gt; 1460

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 54

acgcgtccgc	gcagaaccag	gaaagtaacg	gctacagaca	gtgagaaata	gtttcgctcg	60
ccggctagaa	aaactctgtc	ggtaccaacc	ccagagcgtt	gagagcagcc	cacctccacg	120
cttccttaac	ggagaggtgc	aggactcaga	cttcaccagc	ccactcggtc	ccagccttgt	180
acgcaaagag	acgtcaagga	cgcgctctcc	cgcgtccagg	cagccccagc	ttgctggctt	240
gcctgccgcg	ctgcgtgcag	cactcggccg	gcgtgcagca	tgacctgtg	gaacggcgta	300
ctgccttttt	acccccagcc	ccggcatgcc	gcaggcttca	gcgttccact	gctcatcggt	360
attctagtgt	ttttggctct	agcagcaagc	ttcctgctca	tcttgccggg	gatccgtggc	420
cactcgcgct	ggttttggtt	ggtgagagtt	cttctcagtc	tgttcatagg	cgcagaaatt	480
gtggctgtgc	acttcagtgc	agaatggttc	gtgggtacag	tgaacaccaa	cacatcctac	540
aaagccttca	gcgcagcgcg	cgttacagcc	cgtgtcggtc	tgctcgtggg	cctggagggc	600
attaatatta	cactcacagg	gaccccagtg	catcagctga	acgagaccat	tgactacaac	660
gagcagttca	cctggcgtct	gaaagagaat	tacgccgcgg	agtacgcgaa	cgcactggag	720
aaggggctgc	cggaccagtg	gctctacctg	gcggagaagt	tcacaccgag	tagcccttgc	780
ggcctgtacc	accagtacca	cctggcggga	cactacgcct	cggccacgct	atgggtggcg	840
ttctgcttct	ggctcctctc	caacgtgctg	ctctccacgc	cggccccgct	ctacggaggc	900
ctggcactgc	tgaccaccgg	agccttcgcg	ctcttcgggg	tcttcgcctt	ggcctccatc	960
tctagcgtgc	cgtcttgccc	gctccgccta	ggctcctccg	cgtctaccac	tcagtacggc	1020
gccgccttct	gggtcacgct	ggcaaccggc	gtcctgtgcc	tcttcctccc	aggggcccgtg	1080
gtgagtctcc	agtatgttcg	gcccagcgtc	cttcgcaccc	ttctggacca	aagcgccaag	1140
gactgcaagc	cagaagaggg	ggggctcacc	tcttatcctg	gggaccact	gcacaagcag	1200
gccgctttcc	cagactttaa	aatgtatcac	cactaacctg	tgagggggac	ccaatctgga	1260
ctccttcccc	gccttgggac	atcgcaggcc	gggaagcagt	gcccgcagag	cctggggccag	1320
gagagctcca	ggaagggcac	tgagcgtgcg	tggcgcgagg	cctcggacat	ccgcaggcac	1380
cagggaaagt	ctcctggggc	gatctgtaaa	taaacctttt	tttcttttgt	tttttaaaaa	1440
aaaaaaaaaa	aaaaaaaaaa					1460

&lt;210&gt; 55

&lt;211&gt; 1010

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 55

ccacgcgtcc	gaaattctta	accatgggtc	agcttatatc	gaccttttgg	gtaaaagaca	60
ttattccttt	gaaacttcag	aagcacgagt	ttgttttctg	aaaaacacat	gccattgcat	120
accttttaac	ttgacatatg	actgaaactt	ttgtcttaac	aatccatttt	ctacaaagac	180
tgcccttata	tttatttaaag	atgacaaaac	caaaaccact	tcttctttca	tacattttct	240
gctttcatgt	tctgtcatte	tgagatcttt	ttttctttcc	ctctctctgt	cttcttttcc	300
cccagctgtc	ggcagcatcc	agccccctca	gtcacagtcc	tcacagagct	tcaggaaagg	360
acctctttgc	agagctctct	ttggaggatt	tcttataaat	cacttttaggt	attgctactt	420
gttgtgggag	atggggactt	aaatacttaa	aaacaattac	caaacattaa	acacaattca	480
cgcttctctt	tcatttcttt	aaaggactct	aggctccctc	tcctcctcct	cctcttcctc	540
tattctttag	gtgcagaatg	aaaagatagg	acataaggaa	aaacaacttt	tggaaaatat	600

ctgtaacata	attattttaa	gtaatctcat	ccactgctca	cagtaaattgg	atgagatcgg	660
ataaagtttg	ctttttaatt	aagggcaagt	gaataacttaa	gccctttcga	ttagaaggat	720
tctgcatatg	atccacactg	cctctgtttg	ctggaagaca	gtgtaatttg	ttcttcttga	780
tgctctagcc	cccagtttag	atataccttc	tagatctttg	ggagtttata	gtcttctgat	840
ttggcaagtt	tattagcaca	tctacttcag	taatagatct	ttggatgggc	agtttaaggg	900
aaagtgtttt	tgtatataat	gtattttttc	acttttggag	gattcttttt	gtataacttc	960
aataaagatt	gtaagcaaag	gttgaaaaaa	aaaaaaaaaa	aaaaaaaaaa		1010

&lt;210&gt; 56

&lt;211&gt; 2927

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 56

gctttttcat	aatctttcca	gagtcattct	agtgggattt	ggggaagcaa	cagggctgtg	60
tggggtaacc	tgccaccttt	aagtggark	cagaaatgga	gcaagagcca	caaatggag	120
aacctgctga	aattaagatc	atcagagaag	catataagaa	ggccttttta	tttgtaaca	180
aaggtctgaa	tacagatgaa	ttaggtcaga	aggaagaagc	aaagaactac	tataagcaag	240
gaataggaca	cctgctcaga	gggatcagca	tttcatcaaa	agagtctgaa	cacacaggtc	300
ctgggtggga	atctgctaga	cagatgcaac	agaaaatgaa	agaaactcta	cagaatgtac	360
gcaccaggct	ggaaattcta	gagaagggtc	ttgccacttc	tctgcagaat	gatcttcagg	420
aggtgcccaa	gttatatcca	gaatttccac	ctaaagacat	gtgtgaaaaa	ttaccagagc	480
ctcagtcttt	tagttcagct	cctcagcatg	ctgaagtaaa	tggaaacacc	tcaactccaa	540
gtgcaggggc	agttgctgca	cctgcttctc	tgtctttacc	atcacaaagt	tgtccagcag	600
aagctcctcc	tgcttatact	cctcaagctg	ctgaagggtca	ctacactgta	tcctatggaa	660
cagattctgg	ggagttttca	tcagttggag	aggagtttta	taggaatcat	tctcagccac	720
cgcctcttga	gaccttaggg	ctggatgcag	atgaattgat	tttgatacca	aatggagtac	780
agattttttt	tgtaaatcct	gcaggggagg	ttagtgcacc	ttcgtatcct	gggtaccttc	840
gaattgtgag	gtttttggat	aattctctcg	atacggttct	aaaccgtcct	cccgggtttc	900
ttcaggtttg	tgactggtta	tatcctctag	ttcctgatag	atctccgggt	ctgaaatgta	960
ctgcgggagc	ctacatgttt	cctgatacaa	tgctacaagc	agcaggatgc	tttgtggggg	1020
tcgtcctgtc	ctctgagtta	ccagaggatg	atagagagct	ctttgaggat	ctgttaaggc	1080
aaatgtctga	ccttcggctc	caggccaact	ggaacagagc	agaagaagaa	aatgaattcc	1140
aaatccctgg	aagaactaga	ccctcctctg	accaactaaa	agaagcctct	ggcactgatg	1200
tgaaacagtt	ggaccaaggc	aataaggatg	tacgtcataa	aggaaaacgt	ggaaaaaggg	1260
ctaaagatac	ttcaagtga	gaagttaacc	tgagtcacat	tgtaccatgt	gagccagttc	1320
cagaagaaaa	gccaaaagaa	ttacctgaat	ggagtgaaaa	agtggctcac	aacattttgt	1380
caggtgcttc	ctgggtgagt	tggggtttag	tcaaagggtg	tgagattact	ggtaaggcaa	1440
tccagaaagg	tgcttctaaa	ctccgagagc	ggattcaacc	agaagaaaaa	cccgtggaag	1500
ttagtccagc	tgtcaccaag	ggactttata	tagcgaagca	agctacagga	ggagcagcaa	1560
aagtcagtca	gttcctggtt	gatggagttt	gcactgtagc	aaattgcgtt	ggaaaagaac	1620
tagtccaca	tgtcaagaag	catggaagca	aacttggttc	agaatctctt	aaaaaagaca	1680
aagatgggaa	atctcctctg	gatggtgcta	tggttgtagc	agcaagtagt	gttcaaggat	1740
tttcaactgt	ctggcaagga	ttggaatgtg	cagctaaatg	catcgttaac	aatgtttcag	1800
cagaaactgt	acaaactgtc	agatacaaat	acggatataa	tgcaggagaa	gctaccacc	1860
atgcggtgga	ttctgcggtc	aatgttggcg	taactgccta	caatattaac	aacattggta	1920
tcaaagcaat	ggtgaagaaa	actgcaacac	aaacaggaca	cactctcctt	gaggactatc	1980
agatagttag	taattctcag	agggaaaatc	aagaaggagc	agcaaatgtc	aacgtgagag	2040
gggagaagga	tgagcagacg	aagggaagta	aggaggcaaa	gaagaaagat	aatgatgaa	2100
gtgctgggaa	tcacttatac	caaagcctta	tgaaatggat	gaaattttgt	taaataggca	2160
aatgtggaat	tcctcacaga	ttaaccagta	ttttttaaat	gtattcatte	ctacaaatta	2220
actttcataa	attttatggc	atgtcttcta	tttaaaagga	aaagaataag	tattcttgca	2280
tctggcctta	gaaatgtgaa	gttatattct	caagtttatt	tttttccaag	tgtagctaaa	2340
atatttttgc	aggtaaaata	aagctgatag	tacatgtgtt	gttcaaacct	tgttaaacct	2400

aatattgaac	tatTTTTata	tctgctgtct	ttcagaaggc	aaataggaaa	ctatatattt	2460
gcttaaaaat	tggcatttag	taaccttaat	tctTTTTata	gaaggaatga	cttaaagtat	2520
tgccccctct	ttttgcacta	attgtggatt	tttttagatg	cttctcaaaa	ttttcagtgt	2580
gtaagctaaa	caaaaactaa	aactaagaat	tctcaaaaaa	acttgttcaa	aacagggaaa	2640
gactgatgaa	aagtaaaatg	gactactttt	gtaacttacc	tgtttgttag	gaaatggaat	2700
ggtctctttg	atttaaaatg	aataaaaata	gattattacg	tcttttgtat	tgagactgta	2760
ttgttatgag	cctaggaaat	ttgggaacat	gattgtattg	tattaaaatt	cgaagtgatt	2820
attatcagct	taattggatt	aaaaaagtac	ttcaagaaat	taaaaaaaaa	aaaaaaaaaa	2880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaggg	gggggggc		2927

&lt;210&gt; 57

&lt;211&gt; 1648

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 57

ggcgggagct	caggaccggc	gccttctcct	tgcttctggg	ggcgtggcc	ttgctcccgc	60
tgtgcgaaa	agaatccagg	cccttccacg	cgcgtgtggg	tgcgggggccc	ccgaagtgct	120
cgtgggtccc	cgctaggctc	ccgctggggc	aggaaccgga	atcatgggtg	ggaccaccag	180
caccgcgcg	gtcaccttcg	aggcggacga	gaatgagaac	atcaccgtgg	tgaagggcac	240
ccggctttcg	gaaaatgtga	ttgatcgaat	gaaggaatcc	tctccatctg	gttcgaagtc	300
tcagcggtat	tctggtgctt	atggtgcctc	agtttctgat	gaagaattga	aaagaagagt	360
agctgaggag	ctggcattgg	agcaagccaa	gaaagaatcc	gaagatcaga	aacgactaaa	420
gcaagccaaa	gagctggacc	gagagagggc	tgctgccaat	gagcagttaa	ccagagccat	480
ccttcgggag	aggatatgta	gcgaggagga	acgcgctaag	gcaaagcacc	tggttaggca	540
gctggaagag	aaagaccgag	tgctaaagaa	gcaggatgca	ttctacaaag	aacagctggc	600
tagactggag	gagaggagct	cagagttcta	cagagtcacc	actgaacaat	atcagaaagc	660
tgctgaagag	gtggaagcaa	agttcaagcg	atatgagtct	catccagtct	gtgctgatct	720
gcaggccaaa	attcttcagt	gttaccgtga	gaacacccac	cagaccctca	aatgctccgc	780
tctggccacc	cagtatatgc	actgtgtcaa	tcatgccaaa	cagagcatgc	ttgagaaggg	840
aggataaaaa	ctttcagaat	gagcaaaaaca	ccatcaacgt	taattccaga	gatggaacat	900
tttttttcct	agtgagaaaa	caacccattt	gaagagaaga	ccactaatga	gaagaccact	960
aaagagagac	atcaagaatg	gattcagcag	aatcatttca	cgttttgaac	agcagcagtt	1020
tgaagggcc	aagccttgat	cagggatcag	tcattaaagg	acactcttga	gtattagtaa	1080
accctcttat	gatgattaaa	agagaagggc	agccctctcc	accttttggg	actttctatt	1140
caacttgac	tgaccataaa	atgtttctct	tctgaacaag	ccccatcatt	tggtgaacct	1200
ccaccctaac	aaagtaggat	ggggttgggg	gctaaattaa	ttggagtggg	gcgaggagag	1260
agccagaaaa	catagatccg	agggcagcag	tgctgggtgg	agagagccag	aaaacagatc	1320
tggaggcagc	agtgctggat	ggaattgtct	aggctgtggc	atgttggttt	tgtctttctt	1380
ttctcctttg	attatgtaag	agctatttca	ttataactta	ttatggtgat	tatacaggca	1440
agaagacaaa	aaggagagaa	aatgtacctc	ttctactgga	ataatgttta	tgattacaag	1500
tgagataagg	tatTTTTatc	aatatgaagg	caaccttggc	tgataaaacc	tctatagtga	1560
atactcacat	ctttacttca	ctcactatca	ataataaata	tatTTTctga	caaagaaaaa	1620
aaaaaaaaaa	aaaaaaaagg	gcggccgc				1648

&lt;210&gt; 58

&lt;211&gt; 1014

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 58

ccacgcgtcc	gggctgcccc	ccaagtgtcg	tttgttttac	tgtagggtct	cccgccccgc	60
gccccagtg	ttttctgagg	gcggaaatgg	ccaattcggg	cctgcagttg	ctgggcttct	120



aaaaaaaaaa aaaaaa

1036

&lt;210&gt; 60

&lt;211&gt; 1966

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 60

ctgttcctgg	cggagcgggc	tccgctcgtc	ttctctgtct	tagggctggt	gctggccctg	60
cccacgccta	gggctccggc	gcgtcacggg	cctcagctgg	gattcccgcg	cccctcggac	120
ggccacgaga	ctcggacatc	tttccaggaa	cagcgtgagg	aggacagaag	cacccaacag	180
gactgctcaa	gccacctgcg	aacactgctg	ctaccatgcc	caagagaaag	gcaaaaggag	240
atgctaaagg	tgataaagca	aaggtgaagg	atgagccaca	gaggagatca	gctcggttgt	300
ctgctaaacc	agctcctcca	aaaccagagc	ccaggcctaa	aaaggcctct	gcaaagaagg	360
gagagaagct	tcccaaaggg	agaaagggga	aagcagatgc	tggaaaggat	ggraacaacc	420
ctgcaaaaaa	ccgagatgcc	tctacactcc	agtcccagaa	agcgggaaggc	actgggggatg	480
ccaagtgaag	tgtacatttt	tgagagctct	gtacttatag	tgactctact	gtttgaaata	540
ctattttttt	aatcaagtt	ttataaaaagt	gtagaatttt	ggctttttta	agttatgttg	600
ttagcacaca	ggacacttcc	ttgttgtctt	ttgtggaaag	ggcaagtacc	actaataggg	660
tgtatctcag	aaactgaatt	gaaataaggg	aaaataggat	tttctgtcct	ggtttttgaa	720
gattgttctt	gattcccttg	attcccagga	gagattctct	gacattcacg	tgtcagccac	780
tttggcacgg	aagccttaca	gtgtggggaa	ccaaaacttc	gtgtctcctc	tttccccgat	840
gccatcagca	tagacttgac	ttcctttaa	cagagagttt	gatgtggcct	tggcaaccct	900
aaaatcagct	gtgttaggta	acaaaactca	ggctttctgt	tgatgacatc	gagatgggtg	960
cacttaaaag	agccaagatt	cctgttttca	gtttgtggat	tcattcctgt	ggttttactt	1020
tagtccctcc	atgtcaaagt	gggcctgaga	aaagctcata	catgcctcat	gtgaagtgtc	1080
caccctctct	gaaaatcttt	cttgtttcaa	acagcaacga	catatcttgt	taacttttac	1140
ggtgactttt	ggaggagggg	agtttggaag	ttgtaaaatg	ttatagattg	ttgcctattt	1200
cctgctgaaa	gtaaatgttt	ttaaaaagta	tcataataag	ctgaatacaa	attgggtttg	1260
ggggagatcc	tttctacccc	aaagtcataa	atatattctt	tactgccttg	tggaaatttt	1320
atagttttgc	ctttcacatt	tgtcttttag	ctgtcctgaa	ctgatttttg	tgaaaattat	1380
gagttagaaa	ttcaacttaa	catttttcat	attcgaagtg	gttgtctctg	cactattttac	1440
tgaaaagtgc	atcctttccc	cagtgatttg	taacactgcc	tctttcataa	attaaatttt	1500
tgtgtatgtg	tgtgggtctt	ttgatggttt	ctattctgac	tgacatcaat	ttgtctaata	1560
ttgtagcagt	acagtacagt	cctgattatt	gcaacttttag	gaaaagggtc	gataaaaacc	1620
aagatgcctc	cacattttgt	cataattgta	accaattttca	cctctgtctc	cagtatcacc	1680
acaaaattgt	tcttccttgg	agtgtccttg	ctattcttag	ccaactgttc	ctccatatta	1740
cttctagaat	tagaccaaca	atttataaat	caaacaaacc	caagagcatt	gaaattttga	1800
ttggatttgt	attgaattta	tagattaatc	tggtaaacca	tgtcatcttt	acaatgttgt	1860
cttccaatac	atgaatatgg	tacagctctt	catttactta	ggcctttgaa	atatctttca	1920
ataaagttta	taattttctc	caaaaaaaaa	aaaaaaaaaa	aaattc		1966

&lt;210&gt; 61

&lt;211&gt; 79

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (79)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 61

35

Met Gln Phe Leu Lys Asn Val Arg Cys Phe Thr Phe Ser Phe Ser His  
1 5 10 15  
Leu Val Phe Asn Ile Trp Cys Val Phe Tyr Thr Ser Ser Thr Ser Gln  
20 25 30  
Phe Thr Pro Ala Thr Phe Lys Met Leu Ser Ser Arg Met Gln Leu Val  
35 40 45  
Gln Pro Tyr Leu Gly Pro Val Ser Ala Trp Gly Arg Ala Gly Glu Leu  
50 55 60  
Ser Cys Val His Gly Gly Pro Gly Glu Pro Pro Val Leu Cys Xaa  
65 70 75

<210> 62  
<211> 32  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (32)  
<223> Xaa equals stop translation

<400> 62  
Met Ala Pro Leu Arg Arg Leu Thr Val Trp Arg Val Gly Leu Arg Leu  
1 5 10 15  
Gly Pro Ala Gly Ala Ala Gly Gly Pro Gly Glu Ala Ser Ser Val Xaa  
20 25 30

<210> 63  
<211> 24  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (24)  
<223> Xaa equals stop translation

<400> 63  
Met Glu Val Gly Val Ser Leu Leu Cys Arg Ile Trp Asp Trp Ala Asn  
1 5 10 15  
Val Cys Val Ala Ser Leu Asn Xaa  
20

<210> 64  
 <211> 24  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (4)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (15)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (24)  
 <223> Xaa equals stop translation

<400> 64  
 Gly Pro Gly Xaa Glu Val Ile Glu Ser Trp Glu Pro Ser Leu Xaa Leu  
           1                  5                  10                  15

Cys Cys Ser Gly Gly Gly Glu Xaa  
                   20

<210> 65  
 <211> 38  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (38)  
 <223> Xaa equals stop translation

<400> 65  
 Met Tyr Leu Pro Val Cys Tyr Ile Leu Leu Leu Ile His Pro Ser Arg  
           1                  5                  10                  15

Met Gln Ser Trp Leu Thr Tyr Met Ile Arg Tyr Asn Asn Arg Met Ser  
                   20                  25                  30

Trp Lys Tyr Arg Ser Xaa  
                   35

<210> 66  
 <211> 60  
 <212> PRT  
 <213> Homo sapiens

<220>

<221> SITE

<222> (60)

<223> Xaa equals stop translation

<400> 66

Met Pro Gly Pro His Gln Leu Cys Leu Ile Asn Ala Gly Gln Pro Cys  
1 5 10 15

Ser Leu Val Thr Ser Val Gly His Trp Leu Leu Gly Val Phe Ser Ile  
20 25 30

Leu Pro Cys Ser Leu Leu Ala Pro Gly Glu Lys Gly Pro Tyr Arg Ser  
35 40 45

Leu Phe Gln Val Thr Arg Ser Trp Gly Gly Pro Xaa  
50 55 60

<210> 67

<211> 38

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (38)

<223> Xaa equals stop translation

<400> 67

Leu Ile Ser Leu Ala Phe Leu Leu Pro Gly Ile Leu Gly Lys Thr Asn  
1 5 10 15

Gly Phe Ser Arg His Pro Leu Ser Thr Asn Ala Phe Leu Lys Gln Thr  
20 25 30

Gly Thr Phe Leu Leu Xaa  
35

<210> 68

<211> 72

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (72)

<223> Xaa equals stop translation

<400> 68

Gly Arg Leu Ile Trp Gly Ala Leu Trp Pro Val Pro Asp Cys Gly Cys  
1 5 10 15

Thr Ala Pro Glu Gly Gln Asp Gln Pro Pro Ser Leu Val Cys Cys His  
20 25 30



Gly Cys Arg Pro Gly Ser His Ala Glu Val Trp Gln Gly Cys Ser Arg  
35 40 45

Gly Gln Asp Tyr Ala Gly Phe Ser Val Gly Ser Gly Ala Gly Ala Pro  
50 55 60

Ser Leu Glu Glu Pro Arg Ser Xaa  
65 70

<210> 69

<211> 54

<212> PRT

<213> Homo sapiens

<400> 69

Met Val Arg Ala Gln Met Ser Pro Gln Ser Pro Ala Leu Ala Glu Trp  
1 5 10 15

Ile Leu Leu Leu Ala Leu Leu Leu Arg Leu Arg Trp Arg Gln Leu Trp  
20 25 30

Asn Lys Met Phe Leu Ala Pro Lys Lys Lys Lys Lys Lys Lys Lys Lys  
35 40 45

Lys Lys Lys Lys Lys Lys  
50

<210> 70

<211> 29

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (29)

<223> Xaa equals stop translation

<400> 70

Met Gly Arg Arg Ile Arg Trp His Ser Trp Met Thr Leu Trp Ile Phe  
1 5 10 15

Leu Met Lys Ser Ala Leu Val Trp Met Ile Trp Leu Xaa  
20 25

<210> 71

<211> 4

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (4)

<223> Xaa equals stop translation

<400> 71

Leu Leu Gly Xaa

1

<210> 72

<211> 34

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (34)

<223> Xaa equals stop translation

<400> 72

Met Pro Val Val Ser Val His Arg Pro Val Thr Val Ser Thr Cys Val  
1 5 10 15

Pro Ser Pro Gly His Pro Val Lys Arg Ala His Phe Leu Gln Ser Gly  
20 25 30

Ile Xaa

<210> 73

<211> 36

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (36)

<223> Xaa equals stop translation

<400> 73

Met Arg Leu Leu Thr Ser Gly Trp Leu Leu Trp Ala Arg Thr Gly Pro  
1 5 10 15

Ser Trp Gly Ala Arg Ser Arg Ile Ser Leu Ile Trp Arg Cys Met Ala  
20 25 30

Cys Cys Val Xaa

35

<210> 74

<211> 45

<212> PRT

<213> Homo sapiens

40

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (45)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 74

Val	Leu	Ala	Ala	Phe	Val	Leu	Gly	Ala	Ala	Leu	Ala	Ala	Gly	Leu	Gly
1				5				10						15	

Leu	Val	Cys	Ala	His	Ser	Ala	Pro	His	Ala	Pro	Gly	Pro	Pro	Ala	Arg
			20					25					30		

Ala	Ser	Pro	Ser	Gly	Pro	Gln	Pro	Arg	Arg	Ser	Gln	Xaa
		35					40					45

&lt;210&gt; 75

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (7)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 75

Met	Asn	Thr	Tyr	Leu	Arg	Xaa
1				5		

&lt;210&gt; 76

&lt;211&gt; 47

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (47)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 76

Gly	Arg	Arg	Arg	Ala	Gln	Leu	Leu	Trp	His	Ile	Ile	Phe	Leu	Gln	Leu
1				5				10					15		

Tyr	Phe	Trp	Ala	Arg	Trp	Gln	Arg	Glu	Gly	Ser	Val	Met	Pro	Pro	His
			20					25					30		

Arg	His	Ser	Ile	Arg	Arg	Arg	Val	Asp	Ser	Phe	Trp	Met	Leu	Xaa
		35					40					45		

&lt;210&gt; 77

&lt;211&gt; 70

&lt;212&gt; PRT

41

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (70)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 77

Met	Asp	Gly	Gly	Gln	Ala	Val	Pro	Thr	Leu	Leu	Gly	Gly	Ala	Gly	Leu
1				5					10					15	

Gly	Gly	Val	Phe	His	Cys	Val	Pro	Phe	Ser	Pro	Arg	Pro	Val	Pro	Arg
			20					25					30		

Val	Trp	Gly	Ala	Leu	Leu	Arg	Thr	Pro	Pro	Arg	Ala	Pro	Ser	Ser	Arg
		35					40					45			

Leu	Trp	Thr	Pro	Arg	Gln	Glu	Arg	Pro	Ser	Ala	Ala	Lys	Arg	Asn	Lys
	50					55					60				

Ile	Leu	Ala	Ser	Arg	Xaa
65					70

&lt;210&gt; 78

&lt;211&gt; 47

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (47)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 78

Asn	Leu	Phe	Val	Ser	Val	Tyr	Asp	Ser	Phe	Phe	Ser	Leu	Val	Gln	Ser
1				5					10					15	

Ser	Ala	Ser	Pro	Ala	Leu	Ser	Ser	Thr	Leu	Gln	Ser	Pro	Phe	Arg	Glu
			20					25					30		

Ser	Ile	Ser	Arg	Asn	Glu	Thr	Arg	Gln	Ser	Cys	Ser	Ala	Pro	Xaa
		35					40					45		

&lt;210&gt; 79

&lt;211&gt; 36

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 79

Met	Ile	Ser	Val	Trp	Phe	Asp	Val	Leu	Phe	Leu	Ser	Ser	Val	Ser	Ala
1				5					10					15	

Trp	Val	Ser	Val	Cys	Met	Lys	Trp	Asn	Lys	Leu	Pro	Thr	Ala	Lys	Lys
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

42

20

25

30

Lys Lys Lys Lys  
35

<210> 80  
<211> 18  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (18)  
<223> Xaa equals stop translation

<400> 80  
Cys Val Phe Thr Val Asp Trp Ser Leu Lys Ala Pro Leu Gly Glu Asn  
1 5 10 15

Met Xaa

<210> 81  
<211> 24  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (24)  
<223> Xaa equals stop translation

<400> 81  
Gln His Cys Cys Leu Ser Ser Ala Ser Cys Leu Gly Ser Pro Leu Val  
1 5 10 15

Trp Glu Asp Cys Cys Ala Pro Xaa  
20

<210> 82  
<211> 39  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (39)  
<223> Xaa equals stop translation

<400> 82  
Met Gly Val Leu Pro Gly Leu Phe Pro Phe Pro Pro Ser Leu Ala Thr  
1 5 10 15

43

Val Leu Gly Val Trp Gln Leu Phe Gly Phe Phe Phe Phe Leu Gln Leu  
20 25 30

Cys Val Cys Val Phe Val Xaa  
35

<210> 83  
<211> 82  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (82)  
<223> Xaa equals stop translation

<400> 83  
Met Arg Thr Asn Gln Ser Leu Cys Ser Phe Leu Leu Trp Ser Val Pro  
1 5 10 15

Phe His Gln Ala Ala Cys Pro Gln Ala Lys Asp His Pro Leu Glu Pro  
20 25 30

Ser Met His Pro Glu Gly Thr Gln Leu Gln Ser Cys Ser Thr Met Leu  
35 40 45

Gly Pro Arg Gln Leu Ser Ser Glu Lys Gln Pro Leu Leu Pro Pro Arg  
50 55 60

Ser His Leu Lys Ser Ser Pro Met Leu Arg Ala Cys Lys Gly Leu Thr  
65 70 75 80

Ser Xaa

<210> 84  
<211> 18  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (18)  
<223> Xaa equals stop translation

<400> 84  
Met Ser Ser Leu Leu Leu Ile Ile Ile Leu Ala Leu Ser Leu Ala Tyr  
1 5 10 15

Glu Xaa

<210> 85  
<211> 32  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (32)  
<223> Xaa equals stop translation

<400> 85  
Met Gln Gly Gln Gln Thr Phe Pro Val Lys Gly Gln Ile Gly Ser Ile  
1 5 10 15  
Phe Gly Phe Leu Gly Cys Leu Leu Leu Leu Leu Leu Leu Trp Glu Xaa  
20 25 30

<210> 86  
<211> 17  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (17)  
<223> Xaa equals stop translation

<400> 86  
Met Pro Phe Tyr Cys Leu Ser Leu Leu Leu Leu Phe Phe Lys Leu Ser  
1 5 10 15  
Xaa

<210> 87  
<211> 14  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (14)  
<223> Xaa equals stop translation

<400> 87  
Met Cys Ser Cys Ser Trp Ala Ser Ser Cys Thr Val Trp Xaa  
1 5 10

45

<210> 88  
<211> 32  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (32)  
<223> Xaa equals stop translation

<400> 88  
Met Tyr Val Phe Lys Thr Gln Leu Val Thr Cys Asp Cys Glu Ile Asp  
1 5 10 15  
Asn Thr Trp Met His Ser Cys Ser His Arg Ile Gly Asp Met Leu Xaa  
20 25 30

<210> 89  
<211> 58  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (58)  
<223> Xaa equals stop translation

<400> 89  
Glu Pro Leu Ala Asp Phe Cys Phe Leu Tyr Arg Leu Leu Cys Arg Leu  
1 5 10 15  
Thr Ser Arg Leu Tyr Pro Cys Leu Cys Glu Arg Gly Cys Val Val Gly  
20 25 30

Thr Gly Thr Ile Leu Thr Gln Lys Leu Ala Thr Ser His Cys Leu Ala  
35 40 45

Gln Met Trp Ile Ile Val Ile Thr Asn Xaa  
50 55

<210> 90  
<211> 117  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (117)  
<223> Xaa equals stop translation



46

&lt;400&gt; 90

Ser Lys Asn Ile Phe His Ala Ala Arg Thr Pro Ala Thr Leu Phe Val  
1 5 10 15

Val Ile Phe Ile Thr Tyr Val Ile Ala Gly Val Thr Gly Phe Ile Gly  
20 25 30

Leu Asp Ile Ile Ala Ser Leu Cys Asn Met Ile Met Gly Leu Thr Leu  
35 40 45

Ile Thr Leu Cys Thr Trp Ala Tyr Ile Arg Tyr Ser Gly Glu Tyr Arg  
50 55 60

Glu Leu Gly Ala Val Ile Asp Gln Val Ala Ala Ala Leu Trp Asp Gln  
65 70 75 80

Ala Leu Tyr Lys Leu Tyr Ser Ala Ala Ala Thr His Arg His Leu Tyr  
85 90 95

His Gln Ala Phe Pro Thr Pro Lys Ser Glu Ser Thr Glu Gln Ser Glu  
100 105 110

Lys Lys Lys Met Xaa  
115

&lt;210&gt; 91

&lt;211&gt; 17

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (17)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 91

Met Thr Cys Cys Leu Leu Arg Ser Gly Val Pro Val Cys Leu Ser Leu  
1 5 10 15

Xaa

&lt;210&gt; 92

&lt;211&gt; 13

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (13)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 92

47

Met Ile His Arg Leu Gln Lys Leu Tyr Glu Leu Arg Xaa  
1 5 10

&lt;210&gt; 93

&lt;211&gt; 45

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (45)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 93

Met Ala Gly Leu Ser Leu Pro Thr Leu Phe Phe Phe Asn Met Ile Leu  
1 5 10 15

Asp Cys Leu Gly Pro Trp Glu Asn Ser Ile Tyr Ile Leu Lys Ser Val  
20 25 30

Phe Thr Phe Leu Gln Ile Lys Arg Pro Leu Gly Phe Xaa  
35 40 45

&lt;210&gt; 94

&lt;211&gt; 3

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (3)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 94

Met Leu Xaa  
1

&lt;210&gt; 95

&lt;211&gt; 7

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (7)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 95

Met Trp Pro Ala Pro Ser Xaa  
1 5

<210> 96  
 <211> 29  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (29)  
 <223> Xaa equals stop translation

<400> 96  
 Met Tyr Phe Ile Cys Asn Leu Leu Leu Leu Gly His Ile Tyr Val  
     1                    5                    10                    15  
 Tyr Ile Trp Met Ser Val Tyr Ile Tyr Met Tyr Lys Xaa  
                     20                    25

<210> 97  
 <211> 36  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (31)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (33)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<400> 97  
 Met Ser Ser Leu Leu Leu Ile Ile Ile Leu Ala Leu Ser Leu Ala Tyr  
     1                    5                    10                    15  
 Lys Lys Lys Lys Lys Lys Leu Glu Gly Gly Pro Gly Thr His Xaa Pro  
                     20                    25                    30  
 Xaa Gly Gly Val  
                     35

<210> 98  
 <211> 39  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (39)  
 <223> Xaa equals stop translation

<400> 98

49

Met Gly Lys Ser Val Leu Leu Gly Ser Ile Tyr Tyr Leu Leu Leu Ser  
 1 5 10 15

Ser His Leu Cys Lys Ser Ala Ile Ser Thr Lys Met Cys Asp Arg Arg  
 20 25 30

Ser Gln Arg Ile Leu Leu Xaa  
 35

<210> 99

<211> 57

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (57)

<223> Xaa equals stop translation

<400> 99

Met Trp Pro Phe Tyr Leu Arg Arg Val Val Val Gln Ser Ser Leu Gly  
 1 5 10 15

Pro Met Lys Leu Pro Cys Ser Leu Trp Leu Trp Thr Val Val Ala Ser  
 20 25 30

Cys Ile Cys Cys Pro His Gly Gly Val Phe Leu Cys Gly Ser Cys Ser  
 35 40 45

Cys Cys Val Ser Gly Leu Leu Leu Xaa  
 50 55

<210> 100

<211> 53

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (53)

<223> Xaa equals stop translation

<400> 100

Met Tyr Arg Ala Val Ser Leu Pro Gly Ala Met Gly Trp Ala Cys Pro  
 1 5 10 15

Ala Val Met Ala Pro Cys Phe Leu Leu Leu Gly Ala Gly Ser Arg Thr  
 20 25 30

Leu Pro Thr Ile His Ala Val Ala Pro Arg Ala Glu Pro Leu Leu Val  
 35 40 45

Leu Trp Gln Pro Xaa

50

50

&lt;210&gt; 101

&lt;211&gt; 51

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (51)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 101

Glu	Asn	Ser	Phe	Lys	Lys	Asn	Met	Ser	Leu	Gly	Ser	Cys	Val	Val	Leu
1				5					10				15		

Leu	Pro	Gly	Cys	Lys	Leu	Ser	Pro	Pro	Arg	Ala	Ser	His	Tyr	His	Gln
			20				25					30			

Thr	Ser	Ser	Gly	Pro	Gly	Cys	Glu	Ala	Leu	Phe	Leu	Val	Phe	Arg	Met
	35					40						45			

Asp	Leu	Xaa
	50	

&lt;210&gt; 102

&lt;211&gt; 92

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (92)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 102

Met	Gly	Thr	Trp	Ile	Glu	Met	Ser	Ser	Leu	Trp	Pro	Cys	Thr	Trp	Cys
1				5					10				15		

Thr	Glu	Pro	Trp	Arg	Arg	Ser	Pro	Cys	Pro	Pro	Pro	Cys	Pro	Arg	Pro
			20					25					30		

Ser	Ser	His	Pro	Pro	Arg	Glu	Arg	Arg	Leu	Cys	Ser	Leu	Ala	Pro	Ser
	35						40					45			

Pro	Ser	Cys	Leu	Pro	Ala	Pro	His	Gln	Lys	Thr	Ala	Ser	Ala	Pro	Arg
	50					55				60					

Arg	Pro	Thr	Ala	Ala	Ser	Ala	Ala	Ser	Thr	Ala	Gln	Gly	Ala	Cys	Pro
65					70				75						80

Pro	Ser	Thr	Ala	Ser	Ser	Lys	His	Ser	Gln	Gln	Xaa
				85					90		

<210> 103  
<211> 37  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (37)  
<223> Xaa equals stop translation

<400> 103  
Met Val Phe Pro Phe Leu Ser Phe Ala Ser Ser Cys Thr Leu Ser Met  
1 5 10 15

Val Phe Ser Asp Ser Ile Phe Leu Asn Ser Leu Phe Ile Phe Leu Gly  
20 25 30

Asn Phe Leu Leu Xaa  
35

<210> 104  
<211> 27  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (27)  
<223> Xaa equals stop translation

<400> 104  
Cys Gly Thr Ala Tyr Cys Leu Phe Thr Pro Ser Pro Gly Met Pro Gln  
1 5 10 15

Ala Ser Ala Phe His Cys Ser Ser Leu Phe Xaa  
20 25

<210> 105  
<211> 21  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (21)  
<223> Xaa equals stop translation

<400> 105  
Met Thr Lys Pro Lys Pro Leu Leu Leu Ser Tyr Ile Phe Cys Phe His  
1 5 10 15

Val Leu Ser Phe Xaa  
20

<210> 106  
<211> 73  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (73)  
<223> Xaa equals stop translation

<400> 106  
Met Glu Thr Pro Gln Leu Gln Val Gln Gly Gln Leu Leu His Leu Leu  
1 5 10 15  
Leu Cys Leu Tyr His His Lys Val Val Gln Gln Lys Leu Leu Leu Leu  
20 25 30  
Ile Leu Leu Lys Leu Leu Lys Val Thr Thr Leu Tyr Pro Met Glu Gln  
35 40 45  
Ile Leu Gly Ser Phe His Gln Leu Glu Arg Ser Phe Ile Gly Ile Ile  
50 55 60

Leu Ser His Arg Leu Leu Arg Pro Xaa  
65 70

<210> 107  
<211> 26  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (26)  
<223> Xaa equals stop translation

<400> 107  
Met Glu Leu Ser Arg Leu Trp His Val Gly Phe Val Phe Leu Phe Ser  
1 5 10 15  
Phe Asp Tyr Val Arg Ala Ile Ser Leu Xaa  
20 25

<210> 108  
<211> 212  
<212> PRT  
<213> Homo sapiens

<220>

&lt;221&gt; SITE

&lt;222&gt; (212)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 108

Met Ala Asn Ser Gly Leu Gln Leu Leu Gly Phe Ser Met Ala Leu Leu  
1 5 10 15

Gly Trp Val Gly Leu Val Ala Cys Thr Ala Ile Pro Gln Trp Gln Met  
20 25 30

Ser Ser Tyr Ala Gly Asp Asn Ile Ile Thr Ala Gln Ala Met Tyr Lys  
35 40 45

Gly Leu Trp Met Asp Cys Val Thr Gln Ser Thr Gly Met Met Ser Cys  
50 55 60

Lys Met Tyr Asp Ser Val Leu Ala Leu Ser Ala Ala Leu Gln Ala Thr  
65 70 75 80

Arg Ala Leu Met Val Val Ser Leu Val Leu Gly Phe Leu Ala Met Phe  
85 90 95

Val Ala Thr Met Gly Met Lys Cys Thr Arg Cys Gly Gly Asp Asp Lys  
100 105 110

Val Lys Lys Ala Arg Ile Ala Met Gly Gly Gly Ile Ile Phe Ile Val  
115 120 125

Ala Gly Leu Ala Ala Leu Val Ala Cys Ser Trp Tyr Gly His Gln Ile  
130 135 140

Val Thr Asp Phe Tyr Asn Pro Leu Ile Pro Thr Asn Ile Lys Tyr Glu  
145 150 155 160

Phe Gly Pro Ala Ile Phe Ile Gly Trp Ala Gly Ser Ala Leu Val Ile  
165 170 175

Leu Gly Gly Ala Leu Leu Ser Cys Ser Cys Pro Gly Asn Glu Ser Lys  
180 185 190

Ala Gly Tyr Arg Ala Pro Arg Ser Tyr Pro Lys Ser Asn Ser Ser Lys  
195 200 205

Glu Tyr Val Xaa  
210

&lt;210&gt; 109

&lt;211&gt; 40

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE



54

&lt;222&gt; (40)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 109

Met	Ala	Ala	Lys	Ser	Ala	Thr	Tyr	Trp	Gly	Trp	His	Trp	Val	Asp	Trp
1				5					10					15	

Arg	Trp	Trp	Cys	Thr	Ala	Cys	Gly	Val	Leu	Arg	Phe	Trp	Ser	Gly	Cys
			20					25					30		

Arg	Glu	Gly	Ser	Gln	Leu	Cys	Xaa
	35						40

&lt;210&gt; 110

&lt;211&gt; 27

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (27)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 110

Met	Thr	Ser	Arg	Trp	Cys	His	Leu	Lys	Glu	Pro	Arg	Phe	Leu	Phe	Ser
1				5					10					15	

Val	Cys	Gly	Phe	Ile	Leu	Leu	Val	Leu	Leu	Xaa
			20					25		

&lt;210&gt; 111

&lt;211&gt; 412

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 111

Met	Gly	Asp	Gly	Ala	Val	Lys	Gln	Gly	Phe	Leu	Tyr	Leu	Gln	Gln	Gln
1				5					10					15	

Gln	Thr	Phe	Gly	Lys	Lys	Trp	Arg	Arg	Phe	Gly	Ala	Ser	Leu	Tyr	Gly
			20					25					30		

Gly	Ser	Asp	Cys	Ala	Leu	Ala	Arg	Leu	Glu	Leu	Gln	Glu	Gly	Pro	Glu
		35					40					45			

Lys	Pro	Arg	Arg	Cys	Glu	Ala	Ala	Arg	Lys	Val	Ile	Arg	Leu	Ser	Asp
	50					55					60				

Cys	Leu	Arg	Val	Ala	Glu	Ala	Gly	Gly	Glu	Ala	Ser	Ser	Pro	Arg	Asp
65					70				75						80

Thr	Ser	Ala	Phe	Phe	Leu	Glu	Thr	Lys	Glu	Arg	Leu	Tyr	Leu	Leu	Ala
			85						90					95	

Ala	Pro	Ala	Ala	Glu	Arg	Gly	Asp	Trp	Val	Gln	Ala	Ile	Cys	Leu	Leu
		100						105					110		
Ala	Phe	Pro	Gly	Gln	Arg	Lys	Glu	Leu	Ser	Gly	Pro	Glu	Gly	Lys	Gln
		115					120					125			
Ser	Arg	Pro	Cys	Met	Glu	Glu	Asn	Glu	Leu	Tyr	Ser	Ser	Ala	Val	Thr
		130				135					140				
Val	Gly	Pro	His	Lys	Glu	Phe	Pro	Val	Thr	Met	Arg	Pro	Thr	Glu	Ala
145					150					155					160
Ser	Glu	Arg	Cys	His	Arg	Arg	Gly	Ser	Tyr	Thr	Leu	Arg	Ala	Gly	Glu
			165						170					175	
Ser	Ser	Leu	Glu	Leu	Trp	Gly	Gly	Pro	Glu	Pro	Gly	Thr	Gln	Leu	Tyr
		180						185					190		
Asp	Trp	Pro	Tyr	Arg	Phe	Leu	Arg	Arg	Phe	Gly	Arg	Asp	Lys	Val	Thr
		195					200					205			
Phe	Ser	Phe	Glu	Ala	Gly	Arg	Arg	Cys	Val	Ser	Gly	Glu	Gly	Asn	Phe
	210					215					220				
Glu	Phe	Glu	Thr	Arg	Gln	Gly	Asn	Glu	Ile	Phe	Leu	Ala	Leu	Glu	Glu
225					230					235					240
Ala	Ile	Ser	Ala	Gln	Lys	Asn	Ala	Ala	Pro	Ala	Thr	Pro	Gln	Pro	Gln
			245						250					255	
Pro	Ala	Thr	Ile	Pro	Ala	Ser	Leu	Pro	Arg	Pro	Asp	Ser	Pro	Tyr	Ser
		260						265					270		
Arg	Pro	His	Asp	Ser	Leu	Pro	Pro	Pro	Ser	Pro	Thr	Thr	Pro	Val	Pro
		275					280					285			
Ala	Pro	Arg	Pro	Arg	Gly	Gln	Glu	Gly	Glu	Tyr	Ala	Val	Pro	Phe	Asp
	290					295					300				
Ala	Val	Ala	Arg	Ser	Leu	Gly	Lys	Asn	Phe	Arg	Gly	Ile	Leu	Ala	Val
305					310					315					320
Pro	Pro	Gln	Leu	Leu	Ala	Asp	Pro	Leu	Tyr	Asp	Ser	Ile	Glu	Glu	Thr
			325						330					335	
Leu	Pro	Pro	Arg	Pro	Asp	His	Ile	Tyr	Asp	Lys	Pro	Glu	Gly	Val	Ala
			340					345					350		
Ala	Leu	Ser	Leu	Tyr	Asp	Ser	Pro	Gln	Glu	Pro	Arg	Gly	Glu	Ala	Trp
		355					360					365			
Arg	Arg	Gln	Ala	Thr	Arg	Asp	Arg	Asp	Pro	Ala	Gly	Leu	Gln	His	Val
	370					375					380				

Gln Pro Ala Gly Gln Asp Phe Ser Ala Ser Gly Trp Gln Pro Gly Thr  
 385 390 395 400

Glu Tyr Asp Asn Val Val Leu Lys Lys Gly Pro Lys  
 405 410

<210> 112

<211> 405

<212> PRT

<213> Homo sapiens

<400> 112

Met Gly Asp Gly Ala Val Lys Gln Gly Phe Leu Tyr Leu Gln Gln Gln  
 1 5 10 15

Gln Thr Phe Gly Lys Lys Trp Arg Arg Phe Gly Ala Ser Leu Tyr Gly  
 20 25 30

Gly Ser Asp Cys Ala Leu Ala Arg Leu Glu Leu Gln Glu Gly Pro Glu  
 35 40 45

Lys Pro Arg Arg Cys Glu Ala Ala Arg Lys Val Ile Arg Leu Ser Asp  
 50 55 60

Cys Leu Arg Val Ala Glu Ala Gly Gly Glu Ala Ser Ser Pro Arg Asp  
 65 70 75 80

Thr Ser Ala Phe Phe Leu Glu Thr Lys Glu Arg Leu Tyr Leu Leu Ala  
 85 90 95

Ala Pro Ala Ala Glu Arg Gly Asp Trp Val Gln Ala Ile Cys Leu Leu  
 100 105 110

Ala Phe Pro Gly Gln Arg Lys Glu Leu Ser Gly Pro Glu Gly Lys Gln  
 115 120 125

Ser Arg Pro Phe Met Gly Lys Ser Asn Cys Thr Ala Cys Ala Val Thr  
 130 135 140

Val Gly Pro His Lys Glu Phe Ala Val Thr Met Arg Pro Thr Glu Ala  
 145 150 155 160

Ser Glu Arg Cys His Leu Arg Gly Ser Tyr Thr Leu Arg Ala Gly Glu  
 165 170 175

Ser Ala Leu Glu Leu Trp Gly Gly Pro Lys Pro Gly Thr Gln Leu Tyr  
 180 185 190

Asp Trp Pro Tyr Arg Phe Leu Arg Arg Phe Gly Arg Asp Lys Ala Val  
 195 200 205

Ala Ala Ser Leu Glu Arg Ala Thr Leu Ser Ser Lys Pro Gly Lys Ala  
 210 215 220

Met Arg Ser Ser Trp Pro Trp Lys Arg Pro Ser Leu Pro Arg Arg Met  
 225 230 235 240  
 Leu His Pro Leu His Pro Asn Arg Ser Gln Pro Gln Ser Pro Pro Arg  
 245 250 255  
 Cys Pro Gly Leu Ile Ala Pro Thr Leu Gly Arg Met Thr His Cys Arg  
 260 265 270  
 Arg Leu His Pro Pro His Arg Cys Leu Leu His Gly Leu Gly Ala Arg  
 275 280 285  
 Arg Gly Ser Met Pro Cys Pro Ser Met Arg Trp Pro Val Pro Trp Gly  
 290 295 300  
 Arg Thr Ser Gly Ala Ser Trp Gln Ser Leu Leu Ser Ser Trp Pro Thr  
 305 310 315 320  
 Leu Leu Tyr Asp Ser Ile Glu Glu Thr Leu Pro Pro Arg Pro Asp His  
 325 330 335  
 Ile Tyr Asp Glu Pro Glu Gly Val Ala Ala Leu Ser Leu Tyr Asp Ser  
 340 345 350  
 Pro Gln Glu Pro Arg Gly Glu Ala Trp Arg Arg Gln Ala Thr Ala Asp  
 355 360 365  
 Arg Asp Pro Ala Gly Leu Gln His Val Gln Pro Ala Gly Gln Asp Phe  
 370 375 380  
 Ser Ala Ser Gly Trp Gln Pro Gly Thr Glu Tyr Asp Asn Val Val Leu  
 385 390 395 400  
 Lys Lys Gly Pro Lys  
 405

&lt;210&gt; 113

&lt;211&gt; 56

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 113

Glu Thr Gln Ser His Ser Val Thr Arg Leu Glu Cys Ser Gly Thr Ile  
 1 5 10 15  
 Ser Ala His Cys Asn Leu Cys Leu Pro Gly Ser Ser Asn Ser Pro Ala  
 20 25 30  
 Ser Ala Ser Arg Val Ala Gly Thr Ala Gly Thr Cys Arg Arg Ala Gln  
 35 40 45  
 Leu Ile Phe Val Phe Leu Ala Glu  
 50 55

<210> 114  
 <211> 54  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (48)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (52)  
 <223> Xaa equals stop translation

<400> 114  
 Glu Thr Glu Ser Cys Ser Val Ser Glu Cys Ser Gly Met Val Ser Val  
     1                    5                    10                    15  
 His Cys Asn Leu His Leu Leu Gly Ser Ser Asp Ser Pro Ala Ser Val  
                     20                    25                    30  
 Ser Gln Val Ala Gly Ile Thr Gly Thr Cys Gln His Ala Gln Leu Xaa  
             35                    40                    45  
 Thr Ile Phe Xaa Ala Glu  
             50

<210> 115  
 <211> 45  
 <212> PRT  
 <213> Homo sapiens

<400> 115  
 Arg Leu Glu Cys Ser Gly Thr Ile Ser Ala His Cys Asn Leu Cys Leu  
     1                    5                    10                    15  
 Pro Gly Ser Ser Asn Ser Pro Ala Ser Ala Ser Arg Val Ala Gly Thr  
             20                    25                    30  
 Ala Gly Thr Cys Arg Arg Ala Gln Leu Ile Phe Val Phe  
             35                    40                    45

<210> 116  
 <211> 45  
 <212> PRT  
 <213> Homo sapiens

<400> 116  
 Arg Leu Glu Cys Ser Gly Ala Val Thr Ala His Cys Asn Leu Ser Leu  
     1                    5                    10                    15

59

Leu Gly Ser Ser Asp Pro Pro Ala Ser Ala Ser Arg Val Ala Gly Thr  
20 25 30

Thr Gly Thr His His His Ala Trp Leu Ile Phe Leu Tyr  
35 40 45

&lt;210&gt; 117

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 117

Phe Val Phe Leu Ala Glu Met Gly Phe His His Val Gly Arg Asp Gly  
1 5 10 15

Leu Asp Leu

&lt;210&gt; 118

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 118

Phe Val Phe Leu Val Asp Met Gly Phe His His Val Ser Gln Asp Gly  
1 5 10 15

Leu Asp Val

&lt;210&gt; 119

&lt;211&gt; 18

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 119

Glu Cys Ser Gly Thr Ile Ser Ala His Cys Asn Leu Cys Leu Pro Gly  
1 5 10 15

Ser Ser

&lt;210&gt; 120

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 120

Glu Cys Asn Gly Thr Ile Ser Thr His Cys Asn Leu Cys Leu Pro Cys  
1 5 10 15

60

Gly Thr Ser

&lt;210&gt; 121

&lt;211&gt; 186

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 121

Ile	Glu	Ala	Ala	Ala	Thr	Glu	Val	Val	Asn	Leu	Arg	Asp	Asn	Leu	Asn
1				5					10					15	

Glu	Trp	Asp	Asn	Lys	Val	Gly	Asp	Gly	Asp	Cys	Gly	Ser	Thr	Met	Phe
			20					25					30		

Arg	Gly	Ala	Val	Ala	Ile	Leu	Glu	Asp	Met	Lys	Lys	Tyr	Tyr	Pro	Leu
		35					40					45			

Asn	Asp	Pro	Ala	Glu	Thr	Val	Asn	Glu	Ile	Gly	Ala	Ser	Ile	Gly	Arg
	50					55					60				

Val	Met	Gly	Gly	Thr	Ser	Gly	Ile	Leu	Tyr	Ser	Ile	Phe	Cys	Lys	Ala
65					70					75					80

Ala	Tyr	Ala	Lys	Leu	Lys	Glu	Asn	Ala	Glu	Ser	Val	Val	Thr	Ala	Ile
			85						90					95	

His	Trp	Ala	Asp	Ala	Leu	Glu	Ala	Ala	Ile	Ala	Ala	Val	Ser	Lys	Tyr
			100					105						110	

Gly	Gly	Ala	Ser	Ala	Gly	Tyr	Arg	Thr	Leu	Leu	Asp	Ala	Leu	Ile	Pro
		115					120					125			

Ala	Leu	Ser	Ala	Leu	Lys	Glu	Arg	Leu	Asn	Ala	Gly	Asp	Asp	Pro	Ala
	130					135					140				

Asp	Ala	Phe	Ile	Ile	Ser	Ala	Glu	Ala	Ala	Ser	Ala	Gly	Ala	Glu	Ser
145					150					155					160

Thr	Lys	His	Met	Gln	Ala	Gln	Ala	Gly	Arg	Ser	Thr	Tyr	Val	Pro	Gly
			165						170						175

Asp	Ile	Leu	Ala	Ser	Val	Pro	Asp	Pro	Gly
		180						185	

&lt;210&gt; 122

&lt;211&gt; 181

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 122

Leu	Glu	Arg	Val	Cys	Ser	Thr	Leu	Leu	Gly	Leu	Glu	Glu	His	Leu	Asn
1				5					10					15	

Ala Leu Asp Arg Ala Ala Gly Asp Gly Asp Cys Gly Thr Thr His Ser  
                   20                  25                  30  
 Arg Ala Ala Arg Ala Ile Gln Glu Trp Leu Lys Glu Gly Pro Pro Pro  
                   35                  40                  45  
 Ala Ser Pro Ala Gln Leu Leu Ser Lys Leu Ser Val Leu Leu Leu Glu  
                   50                  55                  60  
 Lys Met Gly Gly Ser Ser Gly Ala Leu Tyr Gly Leu Phe Leu Thr Ala  
                   65                  70                  75                  80  
 Ala Ala Gln Pro Leu Lys Ala Lys Thr Ser Leu Pro Ala Trp Ser Ala  
                   85                  90                  95  
 Ala Met Asp Ala Gly Leu Glu Ala Met Gln Lys Tyr Gly Lys Ala Ala  
                   100                  105                  110  
 Pro Gly Asp Arg Thr Met Leu Asp Ser Leu Trp Ala Ala Gly Gln Glu  
                   115                  120                  125  
 Leu Gln Ala Trp Lys Ser Pro Gly Ala Asp Leu Leu Gln Val Leu Thr  
                   130                  135                  140  
 Lys Ala Val Lys Ser Ala Glu Ala Ala Ala Glu Ala Thr Lys Asn Met  
                   145                  150                  155                  160  
 Glu Ala Gly Ala Gly Arg Ala Ser Tyr Ile Ser Ser Ala Arg Leu Glu  
                   165                  170                  175  
 Gln Pro Asp Pro Gly  
                   180

&lt;210&gt; 123

&lt;211&gt; 310

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 123

Met Phe Trp Arg Arg Leu Asn Glu Ala Ala Val Lys Val Asn Gly Glu  
           1                  5                  10                  15  
 Ala Thr Val Leu Thr Thr His Phe Ser Lys Leu Pro Trp Pro Ser Pro  
                   20                  25                  30  
 Gln Glu Thr Gln Arg Ile Cys Glu Gln Val Arg Ile Ala Ile Glu Glu  
                   35                  40                  45  
 Ile Ile Ile Val Tyr Tyr Ser Leu Pro Lys Asp Gln Gly Ile Thr Leu  
                   50                  55                  60  
 Arg Lys Leu Val Arg Asn Ala Ala Leu Asp Ile Val Asp Gly Met Ala  
                   65                  70                  75                  80



Gln Leu Leu Glu Val Leu Leu Thr Ala Pro Ser Gln Ser Thr Glu Asn  
                     85                    90                    95  
 Gly Asp Leu Ile Ser Cys Asn Ser Val Ser Val Ala Cys Gln Gln Val  
                     100                    105                    110  
 Pro Glu Ile Pro Lys Asp Asn Lys Ala Ala Ala Leu Leu Met Leu Thr  
                     115                    120                    125  
 Lys Ser Val Asp Phe Val Lys Asp Ala His Glu Glu Met Glu Gln Ala  
                     130                    135                    140  
 Val Glu Glu Cys Asp Pro Tyr Ser Gly Leu Leu Asn Asp Ser Glu Asp  
                     145                    150                    155                    160  
 Asn Ser Asp Ser His Ser Asp Glu Asp Gly Val Leu Gly Leu Pro Ser  
                     165                    170                    175  
 Asn Arg Asp Ser Tyr Trp Ser Glu Glu Asp Gln Glu Leu Ile Ile Pro  
                     180                    185                    190  
 Cys Leu Ala Leu Val Arg Ala Ser Arg Ala Ser Leu Lys Lys Ile Arg  
                     195                    200                    205  
 Ile Leu Val Ala Glu Asn Gly Lys Lys Asp Glu Val Ala Gln Leu Asp  
                     210                    215                    220  
 Asp Ile Val Gly His Phe Leu Met Arg Ser Ala Pro Val Trp Ile Asp  
                     225                    230                    235                    240  
 Leu Val Leu Ser Ile Tyr Pro Pro Val Cys His Leu Thr Val Arg Ile  
                     245                    250                    255  
 Ser Ser Ala Lys Leu Val Ser Val Leu Ile Lys Ala Leu Glu Ile Thr  
                     260                    265                    270  
 Lys Ala Ser His Val Ser Pro His Pro Gly Asp Ser Trp Ile Pro Leu  
                     275                    280                    285  
 Leu Ile Asn Ala Val Asp His Cys Met Asn Arg Ile Lys Ala Leu Thr  
                     290                    295                    300  
 Gln Arg Ala Ala Glu Leu  
                     305                    310

&lt;210&gt; 124

&lt;211&gt; 310

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 124

Met Phe Trp Arg Arg Leu Asn Glu Ala Ala Val Thr Val Ser Arg Glu  
           1                    5                    10                    15

Ala Thr Thr Leu Thr Ile Val Phe Ser Gln Leu Pro Leu Pro Ser Pro  
 20 25 30  
 Gln Glu Thr Gln Lys Phe Cys Glu Gln Val His Ala Ala Ile Lys Ala  
 35 40 45  
 Phe Ile Ala Val Tyr Tyr Leu Leu Pro Lys Asp Gln Gly Ile Thr Leu  
 50 55 60  
 Arg Lys Leu Val Arg Gly Ala Thr Leu Asp Ile Val Asp Gly Met Ala  
 65 70 75 80  
 Gln Leu Met Glu Val Leu Ser Val Thr Pro Thr Gln Ser Pro Glu Asn  
 85 90 95  
 Asn Asp Leu Ile Ser Tyr Asn Ser Val Trp Val Ala Cys Gln Gln Met  
 100 105 110  
 Pro Gln Ile Pro Arg Asp Asn Lys Ala Ala Ala Leu Leu Met Leu Thr  
 115 120 125  
 Lys Asn Val Asp Phe Val Lys Asp Ala His Glu Glu Met Glu Gln Ala  
 130 135 140  
 Val Glu Glu Cys Asp Pro Tyr Ser Gly Leu Leu Asn Asp Thr Glu Glu  
 145 150 155 160  
 Asn Asn Ser Asp Asn His Asn His Glu Asp Asp Val Leu Gly Phe Pro  
 165 170 175  
 Ser Asn Gln Asp Leu Tyr Trp Ser Glu Asp Asp Gln Glu Leu Ile Ile  
 180 185 190  
 Pro Cys Leu Ala Leu Val Arg Ala Ser Lys Ala Cys Leu Lys Lys Ile  
 195 200 205  
 Arg Met Leu Val Ala Glu Asn Gly Lys Lys Asp Gln Val Ala Gln Leu  
 210 215 220  
 Asp Asp Ile Val Asp Ile Ser Asp Glu Ile Ser Pro Ser Val Asp Asp  
 225 230 235 240  
 Leu Ala Leu Ser Ile Tyr Pro Pro Met Cys His Leu Thr Val Arg Ile  
 245 250 255  
 Asn Ser Ala Lys Leu Val Ser Val Leu Lys Lys Ala Leu Glu Ile Thr  
 260 265 270  
 Lys Ala Ser His Val Thr Pro Gln Pro Glu Asp Ser Trp Ile Pro Leu  
 275 280 285  
 Leu Ile Asn Ala Ile Asp His Cys Met Asn Arg Ile Lys Glu Leu Thr  
 290 295 300

Gln Ser Glu Leu Glu Leu  
305 310

<210> 125

<211> 265

<212> PRT

<213> Homo sapiens

<400> 125

Leu Tyr Gln Tyr Glu Thr Cys Pro Phe Cys Cys Lys Val Arg Ala Phe  
1 5 10 15

Leu Asp Tyr His Gly Phe Ser Tyr Glu Val Val Glu Val Asn Pro Val  
20 25 30

Thr Arg Ser Gln Ile Lys Phe Ser Thr Thr Tyr Lys Lys Val Pro Ile  
35 40 45

Leu Arg Ser Gly Glu Thr Thr Met Thr Glu Ser Thr Leu Ile Ile Ser  
50 55 60

Thr Leu Ala Thr Tyr Leu Gln Arg Pro Asp Gln Ser Leu Asp Gln Ile  
65 70 75 80

Ile Gln Met Tyr Pro Ala Val Asp Ser Thr Asn Glu Lys Gly Lys Pro  
85 90 95

Val Leu Asn Tyr Pro Asn Lys Phe Phe Val Met Lys Gly Lys Val Asp  
100 105 110

Gly Asp Ala Asn Met Ala Ser Ala Arg Glu Glu Arg Glu Trp Arg Glu  
115 120 125

Trp Val Asp Asn Trp Phe Ile His Leu Ile Ser Pro Asn Val Tyr Arg  
130 135 140

Asn Trp Asn Glu Ser Val Glu Thr Phe Arg Trp Phe Glu Gln Val Gly  
145 150 155 160

Asp Trp His Arg Thr Phe Pro Ala Trp Glu Arg Val Leu Ala Val Tyr  
165 170 175

Val Gly Ala Ala Ala Met Phe Leu Leu Ser Lys Thr Leu Lys Lys Lys  
180 185 190

His Asn Ile Asn Asp Glu Arg Glu Glu Leu Arg Lys Ala Cys Arg Asp  
195 200 205

Trp Met Ala Ala Ile Gly Pro Asn Arg Gln Phe Leu Gly Gly Asp Glu  
210 215 220

Pro Asn Leu Ala Asp Leu Ser Leu Tyr Gly Ala Met Asn Ser Phe Tyr  
225 230 235 240

65

Gly Cys Ser Ala Phe Lys Glu Val Ile Leu Glu Glu Lys Ile Ala Glu  
                             245                            250                            255

Trp Trp Arg Lys Met Asp Ala Leu Val  
                             260                            265

&lt;210&gt; 126

&lt;211&gt; 267

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (263)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;400&gt; 126

Leu Tyr Gln Tyr Lys Thr Cys Pro Phe Cys Ser Lys Val Arg Ala Phe  
   1                            5                            10                            15

Leu Asp Phe His Ala Leu Pro Tyr Gln Val Val Glu Val Asn Pro Val  
                             20                            25                            30

Arg Arg Ala Glu Ile Lys Phe Ser Ser Tyr Arg Lys Val Pro Ile Leu  
                             35                            40                            45

Val Ala Gln Glu Gly Glu Ser Ser Gln Gln Leu Asn Asp Ser Ser Val  
   50                            55                            60

Ile Ile Ser Ala Leu Lys Thr Tyr Leu Val Ser Gly Gln Pro Leu Glu  
   65                            70                            75                            80

Glu Ile Ile Thr Tyr Tyr Pro Ala Met Lys Ala Val Asn Glu Gln Gly  
                             85                            90                            95

Lys Glu Val Thr Glu Phe Gly Asn Lys Tyr Trp Leu Met Leu Asn Glu  
                             100                            105                            110

Lys Glu Ala Gln Gln Val Tyr Gly Gly Lys Glu Ala Arg Thr Glu Glu  
                             115                            120                            125

Met Lys Trp Arg Gln Trp Ala Asp Asp Trp Leu Val His Leu Ile Ser  
   130                            135                            140

Pro Asn Val Tyr Arg Thr Pro Thr Glu Ala Leu Ala Ser Phe Asp Tyr  
   145                            150                            155                            160

Ile Val Arg Glu Gly Lys Phe Gly Ala Val Glu Gly Ala Val Ala Lys  
                             165                            170                            175

Tyr Met Gly Ala Ala Ala Met Tyr Leu Ile Ser Lys Arg Leu Lys Ser  
                             180                            185                            190

Arg His Arg Leu Gln Asp Asn Val Arg Glu Asp Leu Tyr Glu Ala Ala

195					200					205					
Asp	Lys	Trp	Val	Ala	Ala	Val	Gly	Lys	Asp	Arg	Pro	Phe	Met	Gly	Gly
210					215					220					
Gln	Lys	Pro	Asn	Leu	Ala	Asp	Leu	Ala	Val	Tyr	Gly	Val	Leu	Arg	Val
225					230					235					240
Met	Glu	Gly	Leu	Asp	Ala	Phe	Asp	Asp	Leu	Met	Gln	His	Thr	His	Ile
245					250					255					
Gln	Pro	Trp	Tyr	Leu	Arg	Xaa	Glu	Arg	Ala	Ile					
260					265										

```
<210> 127
<211> 71
<212> PRT
<213> Homo sapiens
```

```

<400> 127
Glu Glu Ser Leu Gln Ile Leu Asn Val Lys Thr Pro Leu Asn Arg Glu
  1                      5                      10                      15
Glu Val Glu Lys His Tyr Glu His Leu Phe Asn Ile Asn Asp Lys Ser
          20                      25                      30
Lys Gly Gly Thr Leu Tyr Leu Gln Ser Lys Val Phe Arg Ala Lys Glu
          35                      40                      45
Arg Ile Asp Glu Glu Phe Gly Arg Ile Glu Leu Lys Glu Glu Lys Lys
          50                      55                      60
Lys Glu Glu Asn Ala Lys Thr
  65                      70

```

```
<210> 128
<211> 69
<212> PRT
<213> Homo sapiens
```

```

<400> 128
Gln Glu Ala Gln Gln Ile Leu Asn Val Ser Lys Leu Ser Pro Glu Glu
  1                      5                      10                      15
Val Gln Lys Asn Tyr Glu His Leu Phe Lys Val Asn Asp Lys Ser Val
          20                      25                      30
Gly Gly Ser Phe Tyr Leu Gln Ser Lys Val Val Arg Ala Lys Glu Arg
          35                      40                      45
Leu Asp Glu Glu Leu Lys Ile Gln Ala Gln Glu Asp Arg Glu Lys Gly
  50                      55                      60

```

67

Gln Met Pro His Thr  
65

<210> 129  
<211> 128  
<212> PRT  
<213> Homo sapiens

<400> 129  
Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu  
1 5 10 15  
Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp  
20 25 30  
Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys  
35 40 45  
Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu  
50 55 60  
Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Ile Ile Glu Pro  
65 70 75 80  
Ser Leu Arg Gln Leu Ala Gln Lys Tyr Asn Cys Asp Lys Met Ile Cys  
85 90 95  
Arg Lys Cys Tyr Ala Arg Leu His Pro Arg Ala Val Asn Cys Arg Lys  
100 105 110  
Lys Lys Cys Gly His Thr Asn Asn Leu Arg Pro Lys Lys Lys Val Lys  
115 120 125

<210> 130  
<211> 127  
<212> PRT  
<213> Homo sapiens

<220>  
<221> SITE  
<222> (34)  
<223> Xaa equals stop translation

<220>  
<221> SITE  
<222> (74)  
<223> Xaa equals stop translation

<400> 130  
Met Trp Ile Phe Val Lys Thr Leu Thr Gly Gln Thr Ile Thr Leu Glu

68

1	5	10	15												
Val	Glu	Pro	Ser	Asp	Ala	Thr	Glu	Asn	Val	Lys	Ala	Lys	Ile	Gln	Asp
			20					25						30	
Arg	Xaa	Gly	Ile	Leu	Pro	Asp	Gln	Gln	Cys	Leu	Ile	Phe	Glu	Gly	Lys
		35					40					45			
Gln	Leu	Lys	Asp	Gly	His	Thr	Leu	Leu	Asp	Tyr	Asn	Ile	Gln	Lys	Glu
		50				55					60				
Ser	Thr	Leu	Arg	Leu	Met	Leu	Cys	Leu	Xaa	Gly	Gly	Ile	Thr	Glu	Pro
	65				70					75					80
Ser	Leu	Cys	His	Leu	Ala	Gln	Lys	Tyr	Asn	Cys	Lys	Lys	Met	Met	Cys
			85						90					95	
His	Lys	Ser	Cys	Ala	Arg	Leu	Arg	Pro	His	Ala	Val	Asn	Cys	Cys	Lys
			100					105					110		
Lys	Cys	Gly	His	Thr	Asn	Asn	Leu	Arg	Pro	Lys	Lys	Lys	Val	Glu	
		115					120					125			

&lt;210&gt; 131

&lt;211&gt; 91

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (53)

&lt;223&gt; Xaa equals stop translation

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (82)

&lt;223&gt; Xaa equals stop translation

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (89)

&lt;223&gt; Xaa equals stop translation

&lt;400&gt; 131

Pro	Gln	Leu	Asn	Gly	Tyr	Ile	Glu	Lys	Ser	Thr	Pro	Tyr	Glu	Cys	Gly
1				5					10					15	

Phe	Asp	Pro	Ile	Ser	Pro	Ala	Arg	Val	Pro	Phe	Ser	Ile	Lys	Phe	Phe
			20					25					30		

Leu	Val	Ala	Ile	Thr	Phe	Leu	Leu	Phe	Asp	Leu	Glu	Ile	Ala	Leu	Leu
		35				40						45			

Leu	Pro	Leu	Pro	Xaa	Ala	Leu	Gln	Thr	Thr	Asn	Leu	Pro	Leu	Ile	Val
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

69

50                                      55                                      60  
 Met Ser Ser Leu Leu Leu Ile Ile Ile Leu Ala Leu Ser Leu Ala Tyr  
 65                                      70                                      75                                      80  
 Glu Xaa Leu Gln Lys Gly Leu Asp Xaa Thr Glu  
                                     85                                      90  
  
 <210> 132  
 <211> 183  
 <212> PRT  
 <213> Homo sapiens  
  
 <400> 132  
 Ala Gln Lys Asp Gln Gln Lys Asp Ala Glu Val Glu Gly Leu Ser Ala  
 1                                      5                                      10                                      15  
 Thr Thr Leu Leu Pro Lys Leu Ile Pro Ser Gly Ala Gly Arg Glu Trp  
                                     20                                      25                                      30  
 Leu Glu Arg Arg Arg Ala Thr Ile Arg Pro Trp Gly Thr Phe Val Asp  
                                     35                                      40                                      45  
 Gln Gln Arg Phe Ser Arg Pro Arg Asn Val Gly Glu Leu Cys Gln Arg  
                                     50                                      55                                      60  
 Leu Val Arg Asn Val Glu Tyr Tyr Gln Ser Asn Tyr Val Phe Val Phe  
 65                                      70                                      75                                      80  
 Leu Gly Leu Ile Leu Tyr Cys Val Val Thr Ser Pro Met Leu Leu Val  
                                     85                                      90                                      95  
 Ala Leu Ala Val Phe Phe Gly Ala Cys Tyr Ile Leu Tyr Leu Arg Thr  
                                     100                                      105                                      110  
 Leu Gln Ser Lys Leu Val Leu Phe Gly Arg Glu Val Ser Pro Ala His  
                                     115                                      120                                      125  
 Gln Tyr Ala Leu Ala Gly Gly Val Ser Phe Pro Phe Phe Trp Leu Ala  
                                     130                                      135                                      140  
 Gly Ala Gly Ser Ala Val Phe Trp Val Leu Gly Ala Thr Leu Val Leu  
 145                                      150                                      155                                      160  
 Ile Gly Ser His Ala Ala Phe His Gln Ile Glu Pro Ala Asp Gly Glu  
                                     165                                      170                                      175  
 Glu Leu Gln Met Glu Pro Val  
                                     180

<210> 133  
 <211> 183  
 <212> PRT



Ala Gln Lys Asp Gln Gln Lys Asp Ala Glu Ala Glu Gly Leu Ser Gly  
1 5 10 15

Thr Thr Leu Leu Pro Lys Leu Ile Pro Ser Gly Ala Gly Arg Glu Trp  
20 25 30

Leu Glu Arg Arg Arg Ala Thr Ile Arg Pro Trp Ser Thr Phe Val Asp  
35 40 45

Gln Gln Arg Phe Ser Arg Pro Arg Asn Leu Gly Glu Leu Cys Gln Arg  
50 55 60

Leu Val Arg Asn Val Glu Tyr Tyr Gln Ser Asn Tyr Val Phe Val Phe  
65 70 75 80

Leu Gly Leu Ile Leu Tyr Cys Val Val Thr Ser Pro Met Leu Leu Val  
85 90 95

Ala Leu Ala Val Phe Phe Gly Ala Cys Tyr Ile Leu Tyr Leu Arg Thr  
100 105 110

Leu Glu Ser Lys Leu Val Leu Phe Gly Arg Glu Val Ser Pro Ala His  
115 120 125

Gln Tyr Ala Leu Ala Gly Gly Ile Ser Phe Pro Phe Phe Trp Leu Ala  
130 135 140

Gly Ala Gly Ser Ala Val Phe Trp Val Leu Gly Ala Thr Leu Val Val  
145 150 155 160

Ile Gly Ser His Ala Ala Phe His Gln Ile Glu Ala Val Asp Gly Glu  
165 170 175

Glu Leu Gln Met Glu Pro Val  
180

<213> Homo sapiens

Pro Phe Leu Leu Lys Leu Phe Gln Ser Ile Glu Lys Glu Gly Ile Leu  
1 5 10 15

Pro Asn Ser Phe Tyr Glu Ala Ser Ile Ile Leu Ile Pro Lys Pro Gly  
20 25 30

Arg Asp Thr Thr Lys Lys Glu Asn Phe Arg Pro Ile Ser Leu Met Asn  
35 40 45

Ile Asp Ala Lys Ile Leu Asn Lys Ile Leu Ala Asn Arg Ile Gln Gln  
 50 55 60  
 His Ile Lys Lys Leu Ile His His Asp Gln Val Gly Phe Ile Pro Gly  
 65 70 75 80  
 Met Gln Gly Trp Phe Asn Ile Arg Lys Ser Ile Asn Val Ile Gln His  
 85 90 95  
 Ile Asn Arg Ala Lys Asp Lys Asn His Met Ile Ile Ser Ile Asp Ala  
 100 105 110  
 Glu Lys Ala Phe Asp Lys Ile Gln Gln Pro Phe Met Leu Lys Thr Leu  
 115 120 125  
 Asn Lys Leu Gly Ile Asp Gly Met Tyr Phe Lys Ile Ile Arg Ala Ile  
 130 135 140  
 Tyr Asp Lys Pro Thr Ala Asn Ile Ile Leu Asn Gly Gln Lys Leu Glu  
 145 150 155 160  
 Ala Phe Pro Leu Lys Thr Gly Thr Arg Gln Gly Cys Pro Leu Ser Pro  
 165 170 175  
 Leu Leu Phe Asn Ile Val Leu Glu Val Leu Ala Arg Ala Ile Arg Gln  
 180 185 190  
 Glu Lys Glu Ile Lys Gly Ile Gln Leu Gly Lys Glu Glu Val Lys Leu  
 195 200 205  
 Ser Leu Phe Ala Asp Asp Met Ile Val Tyr Leu Glu Asn Pro Ile Val  
 210 215 220  
 Ser Ala Gln Asn Leu Leu Lys Leu Ile Ser Asn Phe Ser Lys Val Ser  
 225 230 235 240  
 Gly Tyr Lys Ile Asn Val  
 245

&lt;210&gt; 135

&lt;211&gt; 232

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (48)

&lt;223&gt; Xaa equals any of the naturally occurring L-amino acids

&lt;220&gt;

&lt;221&gt; SITE

&lt;222&gt; (51)

&lt;223&gt; Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (52)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (80)  
 <223> Xaa equals any of the naturally occurring L-amino acids

<220>  
 <221> SITE  
 <222> (119)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (128)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (151)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (186)  
 <223> Xaa equals stop translation

<220>  
 <221> SITE  
 <222> (196)  
 <223> Xaa equals stop translation

<400> 135  
 Pro Ile Pro Leu Lys Leu Phe Gln Lys Thr Glu Glu Lys Lys Leu Cys  
     1                    5                    10                    15  
 Asn Thr Phe Tyr Glu Ala Ser Ile Thr Trp Ile Pro Lys Pro Asp Lys  
             20                    25                    30  
 Tyr Thr Thr Arg Lys Asp Lys Leu Asp Thr Lys Ile Phe Tyr Leu Xaa  
             35                    40                    45  
 Lys Tyr Xaa Xaa Ala Lys Phe Ser Ser Ile Leu Lys Arg Ile Ile Tyr  
     50                    55                    60  
 His Asp His Met Glu Phe Ile Ser Arg Ile Gln Gly Trp Phe Ser Xaa  
     65                    70                    75                    80  
 Arg Val Ser Ile Asn Val Ile His His Ile Asn Thr Met Lys Asp Lys  
             85                    90                    95  
 Lys Ser Met Gln Ser Phe Gln Phe Thr Gln Asn Lys His Leu Thr Lys

73

100	105	110
Phe Tyr Asn Leu Ser Gly Xaa Lys His Gly Thr Phe Lys Met Cys Xaa		
115	120	125
Arg Gly Phe Val Lys Asn Pro Gln Leu Ile Tyr Leu Met Arg Lys Lys		
130	135	140
Thr Glu Lys Phe Leu Pro Xaa Gln Arg Val Arg Tyr Gly Cys Met Leu		
145	150	155
Tyr Pro Val Leu Leu Asn Asn Val Leu Lys Val Leu Thr Met Thr Ile		
165	170	175
Arg Gln Ile Asn Lys Gln Thr Gly Lys Xaa Ile Ala Ser Lys Leu Glu		
180	185	190
Arg Lys Gln Xaa Asn Ser Leu Phe Asn Asn Asp Ile Phe Tyr Leu Glu		
195	200	205
Asn Gln Arg Ile His Lys Leu Val Glu Leu Ile Asn Lys Ser Ser Lys		
210	215	220
Ile Ala Gly Tyr Met Phe Asn Ile		
225	230	

<210> 136  
 <211> 45  
 <212> PRT  
 <213> Homo sapiens

<400> 136  
 Leu Leu Met Ile Ile Thr Phe Trp Leu Pro Gln Leu Asn Gly Tyr Met  
 1 5 10 15  
 Glu Lys Ser Thr Pro Tyr Glu Cys Gly Phe Asp Pro Met Ser Pro Ala  
 20 25 30  
 Arg Val Pro Phe Ser Met Lys Phe Phe Leu Val Ala Ile  
 35 40 45

<210> 137  
 <211> 45  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> SITE  
 <222> (8)  
 <223> Xaa equals stop translation

<400> 137  
 Leu Leu Ile Ile Ile Thr Phe Xaa Leu Pro Gln Leu Asn Gly Tyr Ile

74

1	5	10	15
Glu Lys Ser Thr Pro Tyr Glu Cys Gly Phe Asp Pro Ile Ser Pro Ala			
	20	25	30
Arg Val Pro Phe Ser Ile Lys Phe Phe Leu Val Ala Ile			
	35	40	45

<210> 138  
 <211> 120  
 <212> PRT  
 <213> Homo sapiens

<400> 138
Met Asp Asp Ser Lys Val Val Gly Gly Lys Val Lys Lys Pro Gly Lys
1 5 10 15
Arg Gly Arg Lys Pro Ala Lys Ile Asp Leu Lys Ala Lys Leu Glu Arg
20 25 30
Ser Arg Gln Ser Ala Arg Glu Cys Arg Ala Arg Lys Lys Leu Arg Tyr
35 40 45
Gln Tyr Leu Glu Glu Leu Val Ser Ser Arg Glu Arg Ala Ile Cys Ala
50 55 60
Leu Arg Glu Glu Leu Glu Met Tyr Lys Gln Trp Cys Met Ala Met Asp
65 70 75 80
Gln Gly Lys Ile Pro Ser Glu Ile Lys Ala Leu Leu Thr Gly Glu Glu
85 90 95
Gln Asn Lys Ser Gln Gln Asn Ser Ser Arg His Thr Lys Ala Gly Lys
100 105 110
Thr Asp Ala Asn Ser Asn Ser Trp
115 120

<210> 139  
 <211> 120  
 <212> PRT  
 <213> Homo sapiens

<400> 139
Met Asp Asp Ser Lys Val Val Gly Gly Lys Val Lys Lys Pro Gly Lys
1 5 10 15
Arg Gly Arg Lys Pro Ala Lys Ile Asp Leu Lys Ala Lys Leu Glu Arg
20 25 30
Ser Arg Gln Ser Ala Arg Glu Cys Arg Ala Arg Lys Lys Leu Arg Tyr
35 40 45

75

Gln Tyr Leu Glu Glu Leu Val Ser Ser Arg Glu Arg Ala Ile Cys Ala  
50 55 60

Leu Arg Glu Glu Leu Glu Met Tyr Lys Gln Trp Cys Met Ala Met Asp  
65 70 75 80

Gln Gly Lys Ile Pro Ser Glu Ile Lys Ala Leu Leu Thr Gly Glu Glu  
85 90 95

Gln Asn Lys Ser Gln Gln Asn Ser Ser Arg His Thr Lys Ala Gly Lys  
100 105 110

Thr Asp Ala Asn Ser Asn Ser Trp  
115 120

&lt;210&gt; 140

&lt;211&gt; 103

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 140

Ser Arg Gly Phe Lys Ser Ile Ser Ser Leu Ala Val Ser Asp Cys Gly  
1 5 10 15

Asn Phe Thr Ala Val Gly Thr Met Ser Gly Ser Val Leu Val Phe Asp  
20 25 30

Thr His Glu Cys Arg Arg Leu Tyr Phe Ser Pro Glu Ser His Gly Leu  
35 40 45

Phe Val Thr Gly Ile Glu Phe Val Ser Arg Thr Ser Pro Ser Ile Cys  
50 55 60

Glu Asp Ile Gln Ser Glu Thr Pro Gly Ile Ala Ser Gly Phe Gln Ser  
65 70 75 80

Ala Val Val Thr Leu Ala Ala Asp Lys Thr Met Gln Leu His Arg Val  
85 90 95

Pro Tyr Pro Gln Pro Gln Pro  
100

&lt;210&gt; 141

&lt;211&gt; 372

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 141

Gln Glu Val Lys Glu Ile Phe Met His Ser Gly Leu Thr Gln Asn Leu  
1 5 10 15

Leu Ala His Ile Trp Ala Leu Ala Asp Thr Arg Gln Thr Gly Lys Leu  
20 25 30

Ser Lys Glu Gln Phe Ala Leu Ala Met Tyr Phe Ile Gln Gln Lys Val  
 35 40 45  
 Ser Lys Gly Ile Asp Pro Pro Gln Val Leu Ser Pro Asp Met Val Pro  
 50 55 60  
 Pro Ser Glu Arg Gly Thr Pro Ile Pro Asp Ser Ser Ser Thr Leu Ala  
 65 70 75 80  
 Ser Gly Glu Phe Thr Gly Val Lys Glu Leu Asp Asp Ile Ser Gln Glu  
 85 90 95  
 Ile Ala Gln Leu Gln Arg Glu Lys Tyr Ser Leu Glu Gln Asp Ile Arg  
 100 105 110  
 Glu Lys Glu Glu Ala Ile Lys Gln Lys Thr Ser Glu Val Gln Glu Leu  
 115 120 125  
 Gln Asn Asp Leu Asp Arg Glu Thr Ser Ser Leu Gln Glu Leu Glu Ala  
 130 135 140  
 Gln Lys Gln Asp Ala Gln Asp Arg Leu Asp Glu Met Asp Gln Gln Lys  
 145 150 155 160  
 Ala Lys Leu Arg Asp Met Leu Ser Asp Val Arg Gln Lys Cys Gln Asp  
 165 170 175  
 Glu Thr Gln Thr Ile Ser Ser Leu Lys Thr Gln Ile Gln Ser Gln Glu  
 180 185 190  
 Ser Asp Leu Lys Ser Gln Glu Asp Asp Leu Asn Arg Ala Lys Ser Glu  
 195 200 205  
 Leu Asn Arg Leu Gln Gln Glu Glu Thr Gln Leu Glu Gln Ser Ile Gln  
 210 215 220  
 Ala Gly Arg Ala Gln Leu Glu Thr Ile Leu Arg Ser Leu Lys Cys Thr  
 225 230 235 240  
 Gln Asp Asp Ile Asn Gln Ala Arg Ser Lys Leu Ser Gln Leu Gln Glu  
 245 250 255  
 Ser His Leu Glu Ala His Arg Ser Leu Glu Gln Tyr Asp Gln Val Pro  
 260 265 270  
 Asp Gly Val Ser Gly Thr Ser Leu Pro Asp Leu Ala Thr Leu Asn Glu  
 275 280 285  
 Gly Ile Leu Leu Ala Glu Arg Gly Gly Phe Gly Ala Met Asp Asp Pro  
 290 295 300  
 Phe Lys Asn Lys Ala Leu Leu Phe Ser Asn Asn Ser Gln Glu Leu His  
 305 310 315 320

77

Pro Asp Pro Phe Gln Ala Glu Asp Pro Phe Lys Ser Asp Pro Phe Lys  
                   325                                  330                                  335

Gly Ala Asp Pro Phe Lys Gly Asp Pro Phe Gln Ser Asp Pro Phe Ser  
                   340                                  345                                  350

Glu Gln Gln Thr Ala Ala Thr Asp Pro Phe Gly Gly Asp Pro Phe Lys  
                   355                                  360                                  365

Glu Ser Asp Pro  
                   370

<210> 142

<211> 371

<212> PRT

<213> Homo sapiens

<220>

<221> SITE

<222> (370)

<223> Xaa equals stop translation

<400> 142

Gln Glu Val Lys Glu Ile Phe Met His Ser Gly Leu Thr Gln Asn Leu  
       1                                  5                                  10                                  15

Leu Ala His Ile Trp Ala Leu Ala Asp Thr Arg Gln Thr Gly Lys Leu  
                   20                                  25                                  30

Ser Lys Asp Gln Phe Ala Leu Ala Met Tyr Phe Ile Gln Gln Lys Val  
                   35                                  40                                  45

Ser Lys Gly Ile Asp Pro Pro Gln Val Leu Ser Pro Asp Met Val Pro  
                   50                                  55                                  60

Pro Ser Glu Arg Gly Thr Pro Gly Pro Asp Ser Ser Gly Ser Leu Gly  
       65                                  70                                  75                                  80

Ser Gly Glu Phe Thr Gly Val Lys Glu Leu Asp Asp Ile Ser Gln Glu  
                   85                                  90                                  95

Ile Ala Gln Leu Gln Arg Glu Lys Tyr Ser Leu Glu Gln Asp Ile Arg  
                   100                                  105                                  110

Glu Lys Glu Glu Ala Ile Arg Gln Lys Thr Ser Glu Val Gln Glu Leu  
                   115                                  120                                  125

Gln Asn Asp Leu Asp Arg Glu Thr Ser Ser Leu Gln Glu Leu Glu Ala  
                   130                                  135                                  140

Gln Lys Gln Asp Ala Gln Asp Arg Leu Asp Glu Met Asp Gln Gln Lys  
       145                                  150                                  155                                  160

Ala Lys Leu Arg Asp Met Leu Ser Asp Val Arg Gln Lys Cys Gln Asp



78

	165		170		175
Glu Thr Gln Met Ile Ser Ser Leu Lys Thr Gln Ile Gln Ser Gln Glu					
	180		185		190
Ser Asp Leu Lys Ser Gln Glu Asp Asp Leu Asn Arg Ala Lys Ser Glu					
	195		200		205
Leu Asn Arg Leu Gln Gln Glu Glu Thr Gln Leu Glu Gln Ser Ile Gln					
	210		215		220
Ala Gly Arg Val Gln Leu Glu Thr Ile Ile Lys Ser Leu Lys Ser Thr					
	225		230		235
Gln Asp Glu Ile Asn Gln Ala Arg Ser Lys Leu Ser Gln Leu His Glu					
	245		250		255
Ser Arg Gln Glu Ala His Arg Ser Leu Glu Gln Tyr Asp Gln Val Leu					
	260		265		270
Asp Gly Ala His Gly Ala Ser Leu Thr Asp Leu Ala Asn Leu Ser Glu					
	275		280		285
Gly Val Ser Leu Ala Glu Arg Gly Ser Phe Gly Ala Met Asp Asp Pro					
	290		295		300
Phe Lys Asn Lys Ala Leu Leu Phe Ser Asn Asn Thr Gln Glu Leu His					
	305		310		315
Pro Asp Pro Phe Gln Thr Glu Asp Pro Phe Lys Ser Asp Pro Phe Lys					
	325		330		335
Gly Ala Asp Pro Phe Lys Gly Asp Pro Phe Gln Asn Asp Pro Phe Ala					
	340		345		350
Glu Gln Arg Gln Leu Asn Arg Ser Ile Trp Arg Gly Pro Phe Gln Arg					
	355		360		365
Lys Xaa Pro					
	370				

&lt;210&gt; 143

&lt;211&gt; 253

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 143

Gly Val Arg Arg Lys Arg Val Val Thr Thr Val Thr Tyr Val Leu Met
1 5 10 15

Leu Ala Val Gly Gly Ala Leu Ile Ala Ser Leu Ile Tyr Pro Cys Trp
20 25 30

Ala Ser Gly Ser Gln Met Ile Tyr Thr Gln Phe Arg Gly His Ser Asn

79

35	40	45
Glu Arg Ile Leu Ala Lys Ile Gly Val Glu Ile Gly Leu Gln Lys Val		
50	55	60
Asn Val Thr Leu Lys Phe Glu Arg Leu Leu Ser Ser Asn Asp Val Leu		
65	70	75
Pro Gly Ser Asp Met Thr Glu Leu Tyr Tyr Asn Glu Gly Phe Asp Ile		
	85	90
Ser Gly Ile Ser Ser Met Ala Glu Ala Leu His His Gly Leu Glu Asn		
	100	105
Gly Leu Pro Tyr Pro Met Leu Ser Val Leu Glu Tyr Phe Ser Leu Asn		
	115	120
Gln Asp Ser Phe Asp Trp Gly Arg His Tyr Arg Val Ala Gly His Tyr		
	130	135
Thr His Ala Ala Ile Trp Phe Ala Phe Ala Cys Trp Cys Leu Ser Val		
	145	150
Val Leu Met Leu Phe Leu Pro His Asn Ala Tyr Lys Ser Ile Leu Ala		
	165	170
Thr Gly Ile Ser Cys Leu Ile Ala Cys Leu Val Tyr Leu Leu Leu Ser		
	180	185
Pro Cys Glu Leu Arg Ile Ala Phe Thr Gly Glu Asn Phe Glu Arg Val		
	195	200
Asp Leu Thr Ala Thr Phe Ser Phe Cys Phe Tyr Leu Ile Phe Ala Ile		
	210	215
Gly Ile Leu Cys Val Leu Cys Gly Leu Gly Leu Gly Ile Cys Glu His		
	225	230
Trp Arg Ile Tyr Thr Leu Ser Thr Phe Leu Asp Ala Ser		
	245	250

&lt;210&gt; 144

&lt;211&gt; 242

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 144

Gly Ile Arg Gly His Ser Arg Trp Phe Trp Leu Val Arg Val Leu Leu
1
5
10
15

Ser Leu Phe Ile Gly Ala Glu Ile Val Ala Val His Phe Ser Ala Glu
20
25
30

Trp Phe Val Gly Thr Val Asn Thr Asn Thr Ser Tyr Lys Ala Phe Ser
---

80

35	40	45
Ala Ala Arg Val Thr Ala Arg Val Gly Leu Leu Val Gly Leu Glu Gly		
50	55	60
Ile Asn Ile Thr Leu Thr Gly Thr Pro Val His Gln Leu Asn Glu Thr		
65	70	75 80
Ile Asp Tyr Asn Glu Gln Phe Thr Trp Arg Leu Lys Glu Asn Tyr Ala		
	85	90 95
Ala Glu Tyr Ala Asn Ala Leu Glu Lys Gly Leu Pro Asp Pro Val Leu		
	100	105 110
Tyr Leu Ala Glu Lys Phe Thr Pro Ser Ser Pro Cys Gly Leu Tyr His		
	115	120 125
Gln Tyr His Leu Ala Gly His Tyr Ala Ser Ala Thr Leu Trp Val Ala		
	130	135 140
Phe Cys Phe Trp Leu Leu Ser Asn Val Leu Leu Ser Thr Pro Ala Pro		
145	150	155 160
Leu Tyr Gly Gly Leu Ala Leu Leu Thr Thr Gly Ala Phe Ala Leu Phe		
	165	170 175
Gly Val Phe Ala Leu Ala Ser Ile Ser Ser Val Pro Leu Cys Pro Leu		
	180	185 190
Arg Leu Gly Ser Ser Ala Leu Thr Thr Gln Tyr Gly Ala Ala Phe Trp		
	195	200 205
Val Thr Leu Ala Thr Gly Val Leu Cys Leu Phe Leu Pro Gly Ala Val		
	210	215 220
Val Ser Leu Gln Tyr Val Arg Pro Ser Ala Leu Arg Thr Leu Leu Asp		
225	230	235 240

Gln Ser

&lt;210&gt; 145

&lt;211&gt; 656

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 145

Glu Met Glu Gln Glu Pro Gln Asn Gly Glu Pro Ala Glu Ile Lys Ile
1 5 10 15

Ile Arg Glu Ala Tyr Lys Lys Ala Phe Leu Phe Val Asn Lys Gly Leu
20 25 30

Asn Thr Asp Glu Leu Gly Gln Lys Glu Glu Ala Lys Asn Tyr Tyr Lys

35		40		45
Gln Gly Ile Gly His Leu Leu Arg Gly Ile Ser Ile Ser Ser Lys Glu				
50		55		60
Ser Glu His Thr Gly Pro Gly Trp Glu Ser Ala Arg Gln Met Gln Gln				
65		70		75
Lys Met Lys Glu Thr Leu Gln Asn Val Arg Thr Arg Leu Glu Ile Leu				
		85		90
Glu Lys Gly Leu Ala Thr Ser Leu Gln Asn Asp Leu Gln Glu Val Pro				
		100		105
Lys Leu Tyr Pro Glu Phe Pro Pro Lys Asp Met Cys Glu Lys Leu Pro				
		115		120
Glu Pro Gln Ser Phe Ser Ser Ala Pro Gln His Ala Glu Val Asn Gly				
		130		135
Asn Thr Ser Thr Pro Ser Ala Gly Ala Val Ala Ala Pro Ala Ser Leu				
		145		150
Ser Leu Pro Ser Gln Ser Cys Pro Ala Glu Ala Pro Pro Ala Tyr Thr				
		165		170
Pro Gln Ala Ala Glu Gly His Tyr Thr Val Ser Tyr Gly Thr Asp Ser				
		180		185
Gly Glu Phe Ser Ser Val Gly Glu Glu Phe Tyr Arg Asn His Ser Gln				
		195		200
Pro Pro Pro Leu Glu Thr Leu Gly Leu Asp Ala Asp Glu Leu Ile Leu				
		210		215
Ile Pro Asn Gly Val Gln Ile Phe Phe Val Asn Pro Ala Gly Glu Val				
		225		230
Ser Ala Pro Ser Tyr Pro Gly Tyr Leu Arg Ile Val Arg Phe Leu Asp				
		245		250
Asn Ser Leu Asp Thr Val Leu Asn Arg Pro Pro Gly Phe Leu Gln Val				
		260		265
Cys Asp Trp Leu Tyr Pro Leu Val Pro Asp Arg Ser Pro Val Leu Lys				
		275		280
Cys Thr Ala Gly Ala Tyr Met Phe Pro Asp Thr Met Leu Gln Ala Ala				
		290		295
Gly Cys Phe Val Gly Val Val Leu Ser Ser Glu Leu Pro Glu Asp Asp				
		305		310
Arg Glu Leu Phe Glu Asp Leu Leu Arg Gln Met Ser Asp Leu Arg Leu				
		325		330
				335

Gln Ala Asn Trp Asn Arg Ala Glu Glu Glu Asn Glu Phe Gln Ile Pro  
 340 345 350  
 Gly Arg Thr Arg Pro Ser Ser Asp Gln Leu Lys Glu Ala Ser Gly Thr  
 355 360 365  
 Asp Val Lys Gln Leu Asp Gln Gly Asn Lys Asp Val Arg His Lys Gly  
 370 375 380  
 Lys Arg Gly Lys Arg Ala Lys Asp Thr Ser Ser Glu Glu Val Asn Leu  
 385 390 395 400  
 Ser His Ile Val Pro Cys Glu Pro Val Pro Glu Glu Lys Pro Lys Glu  
 405 410 415  
 Leu Pro Glu Trp Ser Glu Lys Val Ala His Asn Ile Leu Ser Gly Ala  
 420 425 430  
 Ser Trp Val Ser Trp Gly Leu Val Lys Gly Ala Glu Ile Thr Gly Lys  
 435 440 445  
 Ala Ile Gln Lys Gly Ala Ser Lys Leu Arg Glu Arg Ile Gln Pro Glu  
 450 455 460  
 Glu Lys Pro Val Glu Val Ser Pro Ala Val Thr Lys Gly Leu Tyr Ile  
 465 470 475 480  
 Ala Lys Gln Ala Thr Gly Gly Ala Ala Lys Val Ser Gln Phe Leu Val  
 485 490 495  
 Asp Gly Val Cys Thr Val Ala Asn Cys Val Gly Lys Glu Leu Ala Pro  
 500 505 510  
 His Val Lys Lys His Gly Ser Lys Leu Val Pro Glu Ser Leu Lys Lys  
 515 520 525  
 Asp Lys Asp Gly Lys Ser Pro Leu Asp Gly Ala Met Val Val Ala Ala  
 530 535 540  
 Ser Ser Val Gln Gly Phe Ser Thr Val Trp Gln Gly Leu Glu Cys Ala  
 545 550 555 560  
 Ala Lys Cys Ile Val Asn Asn Val Ser Ala Glu Thr Val Gln Thr Val  
 565 570 575  
 Arg Tyr Lys Tyr Gly Tyr Asn Ala Gly Glu Ala Thr His His Ala Val  
 580 585 590  
 Asp Ser Ala Val Asn Val Gly Val Thr Ala Tyr Asn Ile Asn Asn Ile  
 595 600 605  
 Gly Ile Lys Ala Met Val Lys Lys Thr Ala Thr Gln Thr Gly His Thr  
 610 615 620

Leu Leu Glu Asp Tyr Gln Ile Val Asp Asn Ser Gln Arg Glu Asn Gln  
625 630 635 640

Glu Gly Ala Ala Asn Val Asn Val Arg Gly Glu Lys Asp Glu Gln Thr  
645 650 655

<210> 146

<211> 656

<212> PRT

<213> Homo sapiens

<400> 146

Glu Met Glu Gln Glu Pro Gln Asn Gly Glu Pro Ala Glu Ile Lys Ile  
1 5 10 15

Ile Arg Glu Ala Tyr Lys Lys Ala Phe Leu Phe Val Asn Lys Gly Leu  
20 25 30

Asn Thr Asp Glu Leu Gly Gln Lys Glu Glu Ala Lys Asn Tyr Tyr Lys  
35 40 45

Gln Gly Ile Gly His Leu Leu Arg Gly Ile Ser Ile Ser Ser Lys Glu  
50 55 60

Ser Glu His Thr Gly Pro Gly Trp Glu Ser Ala Arg Gln Met Gln Gln  
65 70 75 80

Lys Met Lys Glu Thr Leu Gln Asn Val Arg Thr Arg Leu Glu Ile Leu  
85 90 95

Glu Lys Gly Leu Ala Thr Ser Leu Gln Asn Asp Leu Gln Glu Val Pro  
100 105 110

Lys Leu Tyr Pro Glu Phe Pro Pro Lys Asp Met Cys Glu Lys Leu Pro  
115 120 125

Glu Pro Gln Ser Phe Ser Ser Ala Pro Gln His Ala Glu Val Asn Gly  
130 135 140

Asn Thr Ser Thr Pro Ser Ala Gly Ala Val Ala Ala Pro Ala Ser Leu  
145 150 155 160

Ser Leu Pro Ser Gln Ser Cys Pro Ala Glu Ala Pro Pro Ala Tyr Thr  
165 170 175

Pro Gln Ala Ala Glu Gly His Tyr Thr Val Ser Tyr Gly Thr Asp Ser  
180 185 190

Gly Glu Phe Ser Ser Val Gly Glu Glu Phe Tyr Arg Asn His Ser Gln  
195 200 205

Pro Pro Pro Leu Glu Thr Leu Gly Leu Asp Ala Asp Glu Leu Ile Leu  
 210 215 220  
 Ile Pro Asn Gly Val Gln Ile Phe Phe Val Asn Pro Ala Gly Glu Val  
 225 230 235 240  
 Ser Ala Pro Ser Tyr Pro Gly Tyr Leu Arg Ile Val Arg Phe Leu Asp  
 245 250 255  
 Asn Ser Leu Asp Thr Val Leu Asn Arg Pro Pro Gly Phe Leu Gln Val  
 260 265 270  
 Cys Asp Trp Leu Tyr Pro Leu Val Pro Asp Arg Ser Pro Val Leu Lys  
 275 280 285  
 Cys Thr Ala Gly Ala Tyr Met Phe Pro Asp Thr Met Leu Gln Ala Ala  
 290 295 300  
 Gly Cys Phe Val Gly Val Val Leu Ser Ser Glu Leu Pro Glu Asp Asp  
 305 310 315 320  
 Arg Glu Leu Phe Glu Asp Leu Leu Arg Gln Met Ser Asp Leu Arg Leu  
 325 330 335  
 Gln Ala Asn Trp Asn Arg Ala Glu Glu Glu Asn Glu Phe Gln Ile Pro  
 340 345 350  
 Gly Arg Thr Arg Pro Ser Ser Asp Gln Leu Lys Glu Ala Ser Gly Thr  
 355 360 365  
 Asp Val Lys Gln Leu Asp Gln Gly Asn Lys Asp Val Arg His Lys Gly  
 370 375 380  
 Lys Arg Gly Lys Arg Ala Lys Asp Thr Ser Ser Glu Glu Val Asn Leu  
 385 390 395 400  
 Ser His Ile Val Pro Cys Glu Pro Val Pro Glu Glu Lys Pro Lys Glu  
 405 410 415  
 Leu Pro Glu Trp Ser Glu Lys Val Ala His Asn Ile Leu Ser Gly Ala  
 420 425 430  
 Ser Trp Val Ser Trp Gly Leu Val Lys Gly Ala Glu Ile Thr Gly Lys  
 435 440 445  
 Ala Ile Gln Lys Gly Ala Ser Lys Leu Arg Glu Arg Ile Gln Pro Glu  
 450 455 460  
 Glu Lys Pro Val Glu Val Ser Pro Ala Val Thr Lys Gly Leu Tyr Ile  
 465 470 475 480  
 Ala Lys Gln Ala Thr Gly Gly Ala Ala Lys Val Ser Gln Phe Leu Val  
 485 490 495  
 Asp Gly Val Cys Thr Val Ala Asn Cys Val Gly Lys Glu Leu Ala Pro

85

500					505					510						
His	Val	Lys	Lys	His	Gly	Ser	Lys	Leu	Val	Pro	Glu	Ser	Leu	Lys	Lys	
515					520					525						
Asp	Lys	Asp	Gly	Lys	Ser	Pro	Leu	Asp	Gly	Ala	Met	Val	Val	Ala	Ala	
530					535					540						
Ser	Ser	Val	Gln	Gly	Phe	Ser	Thr	Val	Trp	Gln	Gly	Leu	Glu	Cys	Ala	
545					550					555					560	
Ala	Lys	Cys	Ile	Val	Asn	Asn	Val	Ser	Ala	Glu	Thr	Val	Gln	Thr	Val	
565					570					575						
Arg	Tyr	Lys	Tyr	Gly	Tyr	Asn	Ala	Gly	Glu	Ala	Thr	His	His	Ala	Val	
580					585					590						
Asp	Ser	Ala	Val	Asn	Val	Gly	Val	Thr	Ala	Tyr	Asn	Ile	Asn	Asn	Ile	
595					600					605						
Gly	Ile	Lys	Ala	Met	Val	Lys	Lys	Thr	Ala	Thr	Gln	Thr	Gly	His	Thr	
610					615					620						
Leu	Leu	Glu	Asp	Tyr	Gln	Ile	Val	Asp	Asn	Ser	Gln	Arg	Glu	Asn	Gln	
625					630					635					640	
Glu	Gly	Ala	Ala	Asn	Val	Asn	Val	Arg	Gly	Glu	Lys	Asp	Glu	Gln	Thr	
645					650					655						

&lt;210&gt; 147

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 147

Ala	Ala	Asp	Val	Gln	Arg	Gln	Met	Asn	Gln	Glu	Leu	Ile	Lys	Asn	Asp
1				5					10					15	
Glu	Leu	Trp	Lys	Glu	Arg	Met	Ala	Lys	Leu	Glu	Glu	Asn	Leu	Lys	Lys
			20					25					30		
Thr	Asn	Thr	Ile	Leu	Glu	Lys	Glu	Tyr	Ala	Asn	Ala	Val	Glu	Asn	Val
		35					40					45			
His	Lys	Arg	Phe	Val	Ser	Thr	Ala	Ser	Ser	His	Lys	Val	Pro	Pro	Cys
	50					55					60				
Gln	Asp	Leu	Lys	Ser	Gln	Leu	Leu	Ala	Cys	Tyr	Arg	Ala	His	Pro	Gly
65					70					75					80
Glu	Thr	Leu	Lys	Cys	Met	Glu	Glu	Val	Ala	Gln	Phe	Arg	Gln	Cys	Ile



86

85

90

95

Asp Leu His Arg Val Gln Lys Leu  
100

&lt;210&gt; 148

&lt;211&gt; 103

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 148

Ala Lys His Leu Ala Arg Gln Leu Glu Glu Lys Asp Arg Val Leu Lys  
1 5 10 15

Lys Gln Asp Ala Phe Tyr Lys Glu Gln Leu Ala Arg Leu Glu Glu Arg  
20 25 30

Ser Ser Glu Phe Tyr Arg Val Thr Thr Glu Gln Tyr Gln Lys Ala Ala  
35 40 45

Glu Glu Val Glu Ala Lys Phe Lys Arg Tyr Glu Ser His Pro Val Cys  
50 55 60

Ala Asp Leu Gln Ala Lys Ile Leu Gln Cys Tyr Arg Glu Asn Thr His  
65 70 75 80

Gln Thr Leu Lys Cys Ser Ala Leu Ala Thr Gln Tyr Met His Cys Val  
85 90 95

Asn His Ala Lys Gln Ser Met  
100

&lt;210&gt; 149

&lt;211&gt; 208

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 149

Met Ala Ser Met Gly Leu Gln Val Leu Gly Ile Ser Leu Ala Val Leu  
1 5 10 15

Gly Trp Leu Gly Ile Ile Leu Ser Cys Ala Leu Pro Met Trp Arg Val  
20 25 30

Thr Ala Phe Ile Gly Ser Asn Ile Val Thr Ala Gln Thr Ser Trp Glu  
35 40 45

Gly Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly Gln Met Gln Cys  
50 55 60

Lys Met Tyr Asp Ser Met Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala  
65 70 75 80

87

Arg Ala Leu Met Val Ile Ser Ile Ile Val Gly Ala Leu Gly Met Leu  
                     85                    90                    95  
 Leu Ser Val Val Gly Gly Lys Cys Thr Asn Cys Met Glu Asp Glu Thr  
                     100                    105                    110  
 Val Lys Ala Lys Ile Met Ile Thr Ala Gly Ala Val Phe Ile Val Ala  
                     115                    120                    125  
 Ser Met Leu Ile Met Val Pro Val Ser Trp Thr Ala His Asn Val Ile  
                     130                    135                    140  
 Arg Asp Phe Tyr Asn Pro Met Val Ala Ser Gly Gln Lys Arg Glu Met  
                     145                    150                    155                    160  
 Gly Ala Ser Leu Tyr Val Gly Trp Ala Ala Ser Gly Leu Leu Leu Leu  
                     165                    170                    175  
 Gly Gly Gly Leu Leu Cys Cys Ser Cys Pro Pro Arg Ser Asn Asp Lys  
                     180                    185                    190  
 Pro Tyr Ser Ala Lys Tyr Ser Ala Ala Arg Ser Val Pro Ala Ser Asn  
                     195                    200                    205

&lt;210&gt; 150

&lt;211&gt; 205

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 150

Met Ala Asn Ser Gly Leu Gln Leu Leu Gly Phe Ser Met Ala Leu Leu  
   1                    5                    10                    15  
 Gly Trp Val Gly Leu Val Ala Cys Thr Ala Ile Pro Gln Trp Gln Met  
                     20                    25                    30  
 Ser Ser Tyr Ala Gly Asp Asn Ile Ile Thr Ala Gln Ala Met Tyr Lys  
                     35                    40                    45  
 Gly Leu Trp Met Asp Cys Val Thr Gln Ser Thr Gly Met Met Ser Cys  
                     50                    55                    60  
 Lys Met Tyr Asp Ser Val Leu Ala Leu Ser Ala Ala Leu Gln Ala Thr  
                     65                    70                    75                    80  
 Arg Ala Leu Met Val Val Ser Leu Val Leu Gly Phe Leu Ala Met Phe  
                     85                    90                    95  
 Val Ala Thr Met Gly Met Lys Cys Thr Arg Cys Gly Gly Asp Asp Lys  
                     100                    105                    110

88

Val Lys Lys Ala Arg Ile Ala Met Gly Gly Gly Ile Ile Phe Ile Val  
 115 120 125

Ala Gly Leu Ala Ala Leu Val Ala Cys Ser Trp Tyr Gly His Gln Ile  
 130 135 140

Val Thr Asp Phe Tyr Asn Pro Leu Ile Pro Thr Asn Ile Lys Tyr Glu  
 145 150 155 160

Phe Gly Pro Ala Ile Phe Ile Gly Trp Ala Gly Ser Ala Leu Val Ile  
 165 170 175

Leu Gly Gly Ala Leu Leu Ser Cys Ser Cys Pro Gly Asn Glu Ser Lys  
 180 185 190

Ala Gly Tyr Arg Ala Pro Arg Ser Tyr Pro Lys Ser Asn  
 195 200 205

<210> 151

<211> 71

<212> PRT

<213> Homo sapiens

<400> 151

Glu Pro Gln Arg Arg Ser Ala Arg Leu Ser Ala Lys Pro Ala Pro Pro  
 1 5 10 15

Lys Pro Glu Pro Arg Pro Lys Lys Ala Ser Ala Lys Lys Gly Glu Lys  
 20 25 30

Leu Pro Lys Gly Arg Lys Gly Lys Ala Asp Ala Gly Lys Asp Gly Asn  
 35 40 45

Asn Pro Ala Lys Asn Arg Asp Ala Ser Thr Leu Gln Ser Gln Lys Ala  
 50 55 60

Glu Gly Thr Gly Asp Ala Lys  
 65 70

<210> 152

<211> 71

<212> PRT

<213> Homo sapiens

<400> 152

Glu Pro Gln Arg Arg Ser Ala Arg Leu Ser Ala Lys Pro Ala Pro Pro  
 1 5 10 15

Lys Pro Glu Pro Arg Pro Lys Lys Ala Ser Ala Lys Lys Gly Glu Lys  
 20 25 30

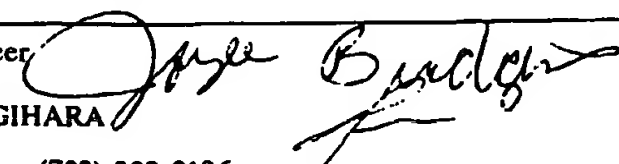
Leu Pro Lys Gly Arg Lys Gly Lys Ala Asp Ala Gly Lys Asp Gly Asn  
 35 40 45

Asn Pro Ala Lys Asn Arg Asp Ala Ser Thr Leu Gln Ser Gln Lys Ala  
50 55 60

Glu Gly Thr Gly Asp Ala Lys  
65 70

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/08983

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> IPC(7) : C07H 21/02, 21/04; C12N 1/20, 15/00, 15/09, 15/63, 15/70, 15/74 US CL : 536/23.1, 23.5, 24.1, 24.2, 24.31, 24.33; 435/252.3, 320.1 According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) U.S. : 536/23.1, 23.5, 24.1, 24.2, 24.31, 24.33; 435/252.3, 320.1  Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Sequence search of public databases.  Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL on EST, AN AI438933. NCI-CGAP (http://www.ncbi.nlm.nih.gov/ncicgap). 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index'. 16 March 1999, see entire sequence.	1-10, 14
X,P	Database EMBL on EST, AN AW130302 NCI-CGAP (http://www.ncbi.nlm.nih.gov/ncicgap). 'National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index'. 28 October 1999, see entire sequence.	1-10, 14
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* "A" "B" "L" "O" "P"	Special categories of cited documents: document defining the general state of the art which is not considered to be of particular relevance earlier document published on or after the international filing date document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *g* document member of the same patent family
Date of the actual completion of the international search 25 MAY 2000		Date of mailing of the international search report 21 AUG 2000.
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer:  NANCY OGIHARA Telephone No. (703) 308-0196

**INTERNATIONAL SEARCH REPORT**International application No.  
PCT/US00/08983**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-10, 14, 18, 21

Remark on Protest

☐  
☐

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

**B. FIELDS SEARCHED**

Electronic data bases consulted (Name of data base and where practicable terms used):

EAST databases: USPatfull, EPO, JPO, Derwent. STN databases: Medline, Caplus, Biosis, Lifesci, Embase. Search terms: secreted proteins, human, cDNA, prostate, substantia nigra, Hodgkins.

**BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING**

This ISA found multiple inventions as follows:

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Group I (claims 1-10, 14, 17, 21), directed to an isolated nucleic acid, whose sequence is selected from SEQ ID NO's 1-60.

Group II (claims 11-12, 16), directed to an isolated polypeptide, where a single sequence is selected from SEQ ID NO's 61-152.

Group III (claim 13), directed an isolated antibody, where a single amino acid sequence to which the antibody is raised is selected from SEQ ID NO's 61-152.

Group IV (claim 15), directed to a method of isolating a polypeptide whose sequence is selected from SEQ ID NO's 61-152.

Group V (claim 18), directed to a method of diagnosis, comprising detecting a polynucleotide whose sequence is selected from SEQ ID NO's 1-60.

Group VI (claim 19), directed to a method of diagnosis, comprising detecting a polypeptide whose amino acid sequence is selected from SEQ ID NO's 61-152.

Group VII (claim 20), directed to a method for identifying a binding partner to a polypeptide, whose amino acid sequence is selected from SEQ ID NO's 61-152.

Group VIII (claim 22), directed to a method for identifying an activity of a polypeptide, whose sequence is selected from SEQ ID NO's 61-152, in a biological assay.

Group IX (claim 23), directed to a binding partner to a polypeptide whose sequence is selected from SEQ ID NO's 61-152.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons:

Group I (claims 1-10, 14, 17, and 21) is drawn to a polynucleotide, whose sequence is recited in SEQ ID NO: 11, which encodes a human secreted protein. The special technical feature of the invention is the sequence recited in SEQ ID NO: 11. The polynucleotide sequence of SEQ ID NO: 11 is not shared by the sequences recited in SEQ ID NOs: 1-10 or 12-60, which encode different polypeptides with different sequences and possess different structural and functional properties. The polynucleotide sequence of SEQ ID NO: 11 is not shared by the amino acid sequences recited in SEQ ID NOs 61-152 which are a completely separate chemical type than polynucleotides. Therefore, there is no shared concept or linking feature between the claims of Group I and Groups II.-VIII (claims 11-13, 15-16, 18-20, 22-23).



Deborah L Pishock/CORP  
08-Nov-2004 13:18

Corporate Intellectual Property UMW22 8276-5979  
To Martine J Denis/RIX/Pharms/SB\_PLC@GSK  
cc  
bcc  
Subject Application title, "Novel Vaccination"  
Our Docket Number: B45307-1

Dear Dr. Denis:

As you are aware, our department filed a continuation-in-part application in the above-identified matter for which you are the sole inventor. I am currently preparing the Information Disclosure Statement (IDS) for this case.

Please send us a list of references (only a list, not references themselves), and any other information which may be material for review by the Patent Examiner. It would be appreciated if you could send us this information as soon as possible, but no later than 01 December 2004. **Please note that failure to disclose a material reference may jeopardize the validity of the resulting patent.**

**What is an IDS?** Each individual associated with the filing and prosecution of a patent application has a duty of candor and good faith in dealing with the U.S. Patent and Trademark Office (PTO). Such duties include disclosure to the PTO all information known to that individual to be material to patentability.

**What is "material?"** Briefly, information is material where there is a substantial likelihood that a reasonable examiner would consider it important in deciding whether to allow the application to issue as a patent. For example, we might have claims to polynucleotides and polypeptides which are more than 80% identical to your gene *over its entire length*. If that is the case, and if you are aware of literature references which disclose polypeptides and polynucleotides which fall within this 80% identity, then these references need to be disclosed. Also, references disclosing proteins or polynucleotides closest to your gene might be material. Note, we claim not only polypeptides and polynucleotides, but antibodies, agonists, antagonists, etc. Please review all claims in the patent application and bring all material references to our attention.

The fact that certain references are cited throughout the body of the specification does not mean that the Patent Office must consider them in determining the patentability of your invention. Therefore, please take care in reviewing the references cited in your specification, such as in the "Background" section, to assess whether you think they may be material and should thus be cited in the IDS that we will soon be filing in the U.S. Patent Office.

You are **not** required to make a literature search, but you are required to disclose relevant references of which you are aware. **Your duty of candor and good faith extends up to the issuance of a patent.**

Examples of references include:

Public database accession numbers;



Journal References;  
Published Patent Applications;  
Issued Patents;  
Related SB Pending Patent Applications (e.g. including any splice  
variants filed by SB); and  
any other information that you believe to be material.

Again, please send the list to me as soon as possible, but no later than  
**01 December 2004**. If you do not have, or are not aware of any references, please send me an  
email stating that fact.

Thank you for your assistance. If you have any questions, please do not hesitate to contact me.

Very truly yours

Deborah L. Pishock  
Assistant to Edward R. Gimmi, Elizabeth J. Hecht and William T. Han  
Corporate Intellectual Property  
GlaxoSmithKline  
709 Swedeland Road, Mailcode UW2220  
P.O. Box 1539  
King of Prussia, PA 19406-0939  
(610)270-5979 (Telephone)  
(610)270-5090 (Fax)

Please note: This e-mail is confidential and may also contain information which is subject to  
Attorney-client privilege and/or subject to attorney work product privilege.

If you are not the intended recipient, you are hereby notified that any dissemination of this  
information is strictly prohibited. If you have received this communication in error please  
notify the sender immediately by reply e-mail and delete the message and any attachments. Thank  
you.